Operational Sea-Ice Classification of Dual Polarized SAR Imagery

by

Mingzhe Jiang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Anthony P. Doulgeris
Professor, Dept. of Physics and Technology
The Arctic University of Norway

Supervisors: David A. Clausi
Professor, Dept. of Systems Design Engineering,
University of Waterloo

Alexander Wong
Professor, Dept. of Systems Design Engineering,
University of Waterloo

Internal Member: Linlin Xu
Research Assistant Professor,
Dept. of Systems Design Engineering,
University of Waterloo

Parsin Haji Reza
Professor, Dept. of Systems Design Engineering,
University of Waterloo

Internal-External Member: Claude Duguay
Professor, Dept. of Geography and Environmental Management,
University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Chapters 2, 3, and 4 contain material from three multi-author papers for which I was the lead author. As the lead author, I made a major contribution on the design, development, evaluation, and writing of the material of papers. The references for the three papers are provided below:

1. Jiang, M., Xu, L. and Clausi, D.A., (2022). "Sea Ice–Water Classification of RADARSAT-2 Imagery Based on Residual Neural Networks (ResNet) with Regional Pooling". *Remote Sensing, 14*(13), p.3025..

2. Jiang, M., Clausi, D.A. and Xu, L., (2022). "Sea Ice Mapping of RADARSAT-2 Imagery by Integrating Spatial Contexture with Textural Features". *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15,* pp.7964-7977.

3. Jiang, M., Xu, L., and Clausi, D., A., (2023). "IceGCN: An Interactive Sea Ice Classification Pipeline for SAR Imagery Based on Graph Convolutional Network". Under review. *IEEE Transactions on Geoscience and Remote Sensing.*

I additionally co-authored five additional papers in related research areas, including:

- Hoekstra, M., Jiang, M. (first co-author), Clausi, D.A. and Duguay, C., (2020). "Lake ice-water classification of RADARSAT-2 images by integrating IRGS Segmentation with pixel-based random forest labeling". *Remote Sensing, 12*(9), p.1425.

- Jiang, M., Chen, X., Xu, L. and Clausi, D.A., 2022, July. "Semi-supervised Sea Ice Classification of SAR Imagery Based on Graph Convolutional Network". In *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1031-1034). IEEE.

- Chen, X., Scott, K.A., Xu, L., Jiang, M., Fang, Y. and Clausi, D.A., (2023). "Uncertainty-Incorporated Ice and Open Water Detection on Dual-Polarized SAR Sea Ice Imagery". *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-13, Art no. 5201213.

- Chen, X., Scott, K. A., Jiang, M., Fang, Y., Xu, L., Clausi, D. A. (2023). "Sea Ice Classification With Dual-Polarized SAR Imagery: A Hierarchical Pipeline". In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 224-232).

- Chen X., Valencia R., Soleymani A., Scott K. A., Jiang M., Xu L., Clausi D. A., (2023) "Uncertainty-incorporated Arctic Sea Ice Concentration Estimation Using Heteroscedastic Bayesian Neural Networks", *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium*, (accepted)

Abstract

Mapping sea ice in polar regions is crucial for research and operational applications, such as environmental modeling and ship navigation. Synthetic aperture radar (SAR) offers a dependable and efficient means of monitoring sea ice under various weather conditions and operational scenarios. Presently, national ice services, such as the Canadian Ice Service (CIS), rely on experienced ice analysts to manually interpret SAR imagery and generate ice charts. Although these charts have been in use for decades, they possess several shortcomings. Manual interpretation necessitates significant expert resources and introduces human bias, and the charts only provide a region-based approximation of ice conditions. Consequently, automated sea-ice classification systems are highly desirable, aiming to accurately label each pixel in SAR imagery with the corresponding sea-ice type.

Addressing the challenges in sea ice classification, such as intra- and inter-class variance, has proven difficult for single-model-based systems due to the limited spatial and contextual information captured. This thesis introduces sea ice classification methods comprising two primary stages: unsupervised segmentation to generate homogeneous regions, and labeling to assign each homogeneous region an appropriate label.

Two innovative methods that directly combine segmentation and labeling for automated sea ice classification are presented. Firstly, a convolutional neural network (CNN) based approach, inspired by the rapid advancements in deep-learning architectures, is developed to differentiate between sea ice and open water. A regional pooling layer is introduced to harness the spatial features learned through labeling and the contextual information extracted via segmentation. Since CNN models necessitate extensive labeled samples, which are scarce for various ice types, a random-forest-based method trained on limited labeled samples is formulated. Texture features are initially extracted from each pixel and then combined with an energy function to assign pixel-level labels of sea ice types to homogeneous regions.

Acknowledging the limitations of direct integration, including the need for extensive labeled training samples and the inherent issue of CNN's limited receptive field, a semi-supervised graph convolutional network (GCN) based method is proposed for operational sea ice classification. A CNN is initially employed to extract features from each node generated by segmentation, followed by the construction of a graph based on the feature vector and the statistical relation between nodes. Lastly, a GCN is designed to propagate local spatial and contextual information to the global level, facilitating the classification of unlabeled nodes in the graph. As a semi-supervised method, the proposed network requires labeled samples from the processing scene as initiating nodes. In comparison to

fully-supervised methods, the GCN-based approach substantially reduces misclassifications of sea ice types due to the absence of prior knowledge.

The classification outcomes reveal that the proposed methods demonstrate promising performance in their respective categories and surpass other state-of-the-art sea-ice classification methods in terms of both numerical accuracy and visual interpretation. Furthermore, the GCN-based method incorporates human supervision to enhance the quality of the produced sea-ice maps, which exhibit high consistency with ice charts released by ice analysts.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, Prof. David Clausi and Prof. Alexander Wong. Their constant guidance, patience, and immense knowledge have been the pillars of support throughout my Ph.D. journey.

I would also like to thank my friends and colleagues in the department, Saeid Taleghanidoozdoozan, Prof. Linlin Xu, Dr. Yuan Fang, Max Manning, Peter Lee, and Javier Noa, for their contributions and support. Special thanks to Dr. Mohsen Ghanbari and Dr. Xinwei Chen for always coming up with great ideas for my research.

Finally, I am eternally grateful to my family for their unconditional love, encouragement, and belief in me, which have been my source of strength and determination throughout this journey.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

1.1.1 Sea ice in SAR imagery

Mapping sea ice in polar regions is crucial not only for a range of research applications, such as oceanography and global climate modeling but also for operational tasks, such as ship navigation. Over the past decades, synthetic aperture radar (SAR) imagery has emerged as the primary data source for sea ice monitoring. As an active microwave remote sensing platform, SAR possesses the unique ability to acquire imagery of targets under any weather conditions, both day and night. The backscattering received by the SAR sensor is largely dependent on the target's electromagnetic properties and surface roughness, both of which significantly distinguish different types of sea ice, inferring its thickness.

The World Meteorological Organisation (WMO) has established a sea ice nomenclature that categorizes sea ice according to several criteria, such as forms of ice, stage of melting, sea ice movement, and sea ice stage of development [1]. Like many related studies, this research will focus on classifying sea ice based on the stage of development. Sea ice's stage of development is usually defined by the thickness of ice formed in less than one year. For any ice that survives through the melting season, the stage of development is described based on survival time. Table 1.1 lists the sea ice stages of development and corresponding thickness, symbols, and color codes. The five main stages can be further subdivided into ten stages. For instance, young ice (10-30 cm) can be subdivided into grey ice (10-15 cm) and grey-white ice (15-30 cm) based on thickness. Any first-year ice that survives through

October 1st becomes old ice. Beyond this stage, the ice is categorized based on the number of survival years, regardless of the thickness.

Table 1.1: Sea-ice stages of development defined by WMO [1], along with corresponding thickness and symbol.

Stage of development	Thickness	Symbol
New Ice	-	1
Nilas	<10 cm	2
Young ice	10-30 cm	3
Grey ice	10-15 cm	4
Grey-white ice	15-30 cm	5
First-year ice	30-200 cm	6
Thin first-year ice	30-70 cm	7
Medium first-year ice	70-120 cm	1•
Thick first-year ice	>120 cm	4•
Old ice	-	7•
Residual ice	-	6•
Second-year ice	-	8•
Multi-year ice	-	9•

The central objective of sea ice mapping using SAR imagery is to delineate the geographic distribution of water and various ice types across vast regions. Presently, the Canadian Ice Service (CIS) interprets between 110 (mid-March) and 290 (early September) SAR images daily to generate sea ice maps, also known as ice charts. Analysts at CIS first segment the entire SAR scene into smaller polygons delineated by distinct boundaries. Each region, which may contain one or multiple ice types, is denoted by the "egg code," which includes the stage of development and other properties as defined by the WMO. An example ice chart is illustrated in Figure 1.1.

Despite their wide usage for several decades, these ice charts have inherent limitations. First, egg codes are denoted for large regions, thereby not identifying ice types at pixel resolution [3]. Second, the generation of polygons is subjective, meaning results can be influenced by human bias, and some variation between analysts may exist. Third, the number of SAR scenes requiring processing is expected to increase over time. Back in 2002, CIS analysts processed around ten scenes daily [4]. Nowadays, with data influx from RADARSAT Constellation Mission (RCM) and Sentinel-1 [5], CIS is interpreting between 110 (mid-March) to 290 (early September) SAR images every day. Therefore, manual

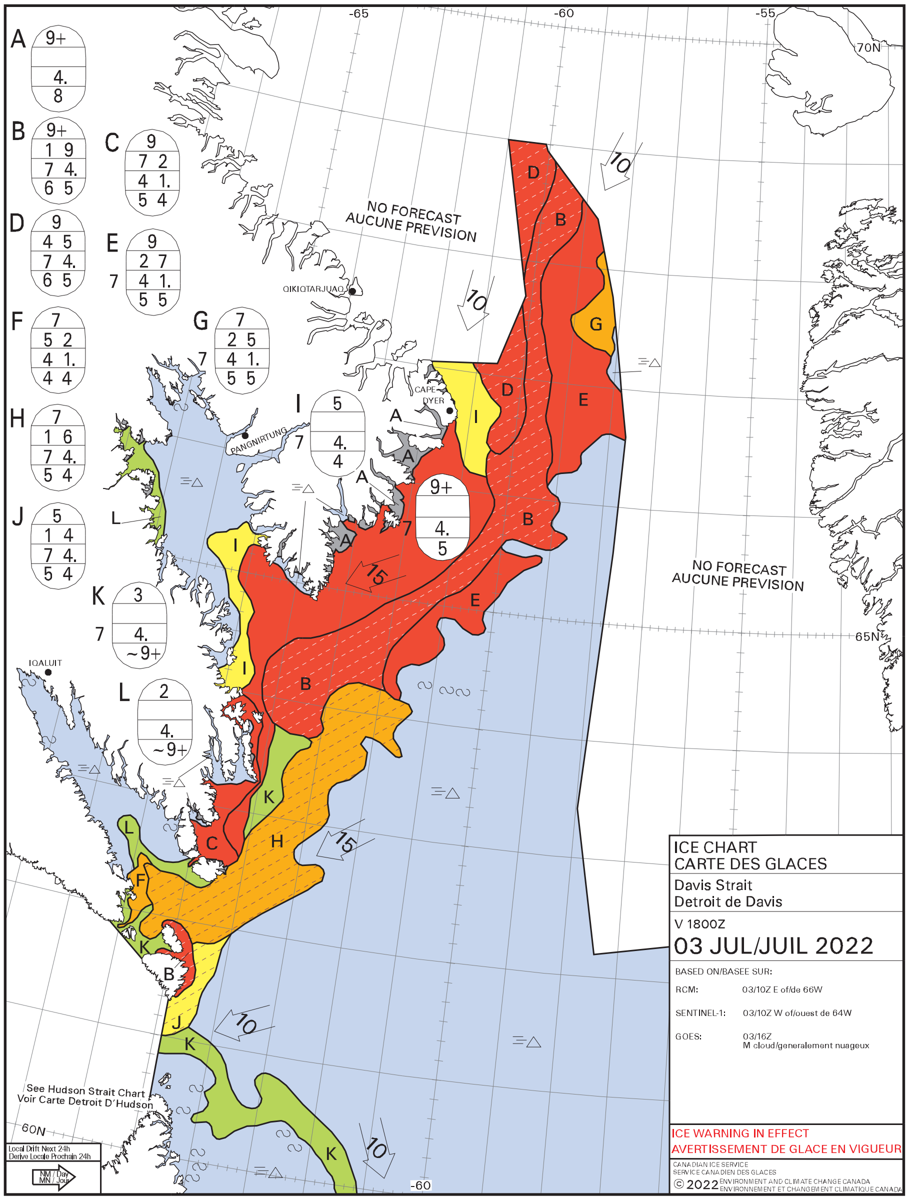


Figure 1.1: An example of regional ice chart of Davis Strait produced by the Canadian Ice Services based on RADARSAT Constellation Mission (RCM) and Sentinel-1 images captured on 03 July 2022. Source: [2]

.

procedures may fall short of future needs.

As a result, there is a pressing need for an automated, computer-based sea ice mapping method. An effective operational automated method should possess the following characteristics:

- The ability to generate ice maps that are consistent with expert interpretation.
- The ability to classify different ice types as required for operational use on full-scene SAR imagery.
- The ability to segment boundaries that match the natural ice and water boundaries.
- The ability to be invariant to SAR sensor artifacts such as banding noise and incidence angle variation.

1.1.2 Challenges

Automated, particularly machine-learning-based, sea ice classification systems hold great promise for national sea ice services, yet their development is fraught with challenges. Typically, the foundation for a classification system rests on modeling the unique distributions of diverse sea ice types, subsequently delineating them within a feature space. A critical determinant of the classification system's performance hinges on the intra-class and inter-class variability of sea ice types. Ideally, a robust machine learning model benefits from low intra-class variability and high inter-class variability, conditions conducive to precise classification results. However, the existence of factors that result in high intra-class variability and low inter-class variability poses significant challenges to accurate sea-ice classification.

Environmental impact

Environmental effects on SAR imagery pose a significant challenge in sea ice classification. A range of environmental factors, such as elevated temperatures leading to the formation of melt ponds on sea ice or strong winds generating waves, can substantially modify the backscatter characteristics. Moreover, the temporal variability of SAR imagery, often a consequence of rapidly changing environmental conditions, can lead to an imbalanced class distribution within the dataset. This dynamic evolution of sea ice conditions can impede the generalizability of the machine learning model, as the model may be trained on conditions that quickly become outdated or rare.

Sensor limitations

In contrast to hyperspectral images that typically comprise over a hundred bands, SAR images used for sea ice monitoring are usually dual-polarized, thereby curtailing the range of available "frequency features", specifically polarization features. Speckle noise, a characteristic of SAR imagery, engenders high-frequency variations in pixel intensity, thus further complicating the classification process. Additionally, the spatial resolution of SAR images used for sea ice monitoring is typically in the order of 100 meters. At this resolution, a single pixel could contain multiple types of sea ice, thereby posing a challenge in assigning a definitive class label to these mixed pixels. These mixed-pixel scenarios are most likely to occur at the boundaries between different sea ice types and water. Misclassification of boundary pixels carries higher implications compared to non-boundary pixels, as boundary information is crucial for operational sea ice mapping. Finally, the incidence angle effect can significantly distort the backscatter characteristics of sea ice, adding another layer of complexity to the classification task. Different sea ice types, such as young ice and first-year ice, exhibit unique backscatter characteristics, which can, however, vary depending on the incidence angle. This means the same type of ice can appear differently, and different types of ice can appear similar in SAR images captured at varying incidence angles.

Machine learning method Limitations

Despite their promise, the employment of machine learning methods, such as convolutional neural networks (CNNs), brings along their own set of challenges. CNNs typically undertake classifications at the pixel level within a rectangular-shaped receptive field. Nevertheless, given the extensive coverage of sea ice, the receptive field may be too confined to effectively capture spatial features critical for accurate classification. Furthermore, the irregular contours of sea ice formations do not neatly fit within the rectangular receptive fields commonly used in CNNs, potentially leading to misclassification. The scarcity of labeled data presents another considerable challenge, as the limited data might be insufficient to train a machine learning model capable of generalizing effectively across diverse sea ice conditions and types.

1.2 Thesis Objectives

The common pipeline for developing sea ice classification methods typically encompasses three stages: data preprocessing, feature extraction, and classifier training. The latter two stages have accumulated significant attention within the research community over the past decade and are the focal points of investigation in this thesis.

Most machine learning methods conventionally rely on spatial features extracted using rectangular filters for feature extraction. Nevertheless, this approach does not consistently yield effective results in the classification of sea ice in SAR images. The quality of spatial features is highly dependent on the dimensions of the rectangular filter. A filter that is too small may fail to capture adequate spatial information, while an excessively large filter could encompass multiple sea ice types, introducing noise and ambiguity. Additionally, factors such as speckle noise and variations in incidence angles can compromise the integrity of spatial features, leading to classification inaccuracies.

On the other hand, semantic segmentation algorithms have demonstrated proficiency in extracting contextual information that is robust to noise and adept at distinguishing features with differing characteristics. However, the types of spatial and contextual features that should be learned, as well as the methods for combining these features, have not been thoroughly investigated, revealing gaps in the research that require further exploration.

Therefore, the primary objective of this thesis is to develop frameworks that inherently integrate spatial and contextual features for the classification of sea ice in SAR images. More specifically, these infrastructures should be capable of generating sea ice maps with pixel-level accuracy and well-defined ice boundaries, leveraging dual-polarization data and a constrained volume of training samples. Crucially, the classification accuracy of the proposed frameworks must be insensitive to temporal changes and incidence angle variations to ensure the reliability of results under varying conditions. The systems should offer functionality that is either fully automated or semi-automated, contingent upon the task at hand—whether it is to distinguish sea ice from open water or to differentiate among various sea ice types.

Moreover, the intended system will be designed to meet the operational requirements of national sea ice services, a benchmark that current models have yet to achieve. In line with this, it will strive to enhance the efficiency and scalability of sea ice mapping operations, ensuring that it can handle an increasing volume of SAR images expected in the future. Furthermore, a special emphasis will be placed on enhancing the system's capability to handle mixed pixels and accurately classify boundary pixels, which often present a significant challenge in sea ice classification.

Ultimately, the thesis seeks to contribute to the broader scientific understanding and practical operational capacity of sea ice classification, with potential ramifications for global climate modeling, oceanography studies, and ship navigation.

1.3 Thesis Structure

The structure of this thesis unfolds over several chapters, each presenting an innovative sea-ice classification technique for dual-pol SAR imagery. These methods have demonstrated encouraging classification results, both qualitatively and quantitatively. Chapters 2, 3, and 4 are grounded in submitted or published manuscripts; thus, there might be some redundancy in the introductory and background sections of these chapters.

Chapter 2 presents IRGS-RF, an automated method for mapping various types of sea ice in RADARSAT-2 imagery by integrating spatial context and textural features. In contrast to the method elucidated in Chapter 2, IRGS-RF does not employ deep learning models due to the scarcity of training data for different sea-ice types. Instead, it proposes an energy function to integrate IRGS and Random Forest (RF).

Chapter 3 introduces IceNet, a robust and automatic approach to distinguishing sea ice from open water in dual-pol SAR imagery. A regional pooling layer is deployed to combine the unsupervised Iterative Region Growing using Semantics (IRGS) segmentation and supervised pixel-wise Residual Network (ResNet) labeling, both of which are state-of-the-art methods in remote sensing. This study marks the first application of a CNN in conjunction with unsupervised segmentation for sea ice-water classification.

Chapter 4 introduces IceGCN, an interactive sea ice classification method for SAR imagery based on a GCN. This innovative approach necessitates human interpretation as prior knowledge to ensure an accurate representation of sea ice types. Differing from the previously mentioned methods, IceGCN accounts for both intra-superpixel and inter-superpixel relations when merging segmentation with labeling.

Chapter 5 provides a summary of the research conducted and contributions made to SAR sea ice classification and discusses potential future work.

The sea-ice classification results included in this thesis adhere to the color code specified by the World Meteorological Organization for different stages of development [6] and are best viewed in color. For ease of reference, the highest numerical accuracies in result tables are highlighted in **bold** typeface.

Chapter 2

Sea Ice Mapping by Integrating Spatial Contexture with Textural Features

Mapping different types of sea ice that form, grow, and melt in polar oceans is essential for shipping navigation, climate change modeling, and local community safety. Currently, ice charts are manually generated by analysts at the Canadian Ice Service (CIS) based on dual-polarized RADARSAT-2/RADARSAT Constellation Mission (RCM) imagery on a daily basis. Inspired by the demand for a computer-based mapping system, we have developed an automatic sea ice classification method that integrates spatial contexture (unsupervised segmentation) with textural features (supervised pixel-level labeling). First, the full-scene image is oversegmented, and the segments are merged into homogeneous regions across the entire scene. Second, pixel-based classifiers (support vector machine, random forest) are compared for their ability to label the generated homogeneous regions. Finally, the segmentation and labeling are combined using a proposed energy function. The proposed method was tested on 18 dual-polarization RADARSAT-2 scenes acquired over the Beaufort Sea. This dataset contains water, young ice, first-year ice, and multi-year ice covering melt, summer, and freeze-up seasons. The proposed method obtains an average classification accuracy of 86.33% based on the leave-one-out validation. The experimental results show that the proposed method achieves promising classification results in both quantity and quality measurements compared to benchmark methods. The robustness against incidence angle variance indicates that the proposed method is well-qualified for operational sea ice mapping.

2.1 Introduction

The interpretation of ice types and analysis of their properties in polar ocean regions have several crucial applications, including ship navigation, global climate monitoring, and animal migration forecasting [7, 8]. For consideration of expense, efficiency, accuracy, and timing requirements, remote sensing has been chosen as an appropriate method for sea ice monitoring. Satellite-based synthetic aperture radar (SAR) is the imaging system of choice for this application since it is not affected by cloud cover and, since it is self-illuminated, can be used equivalently under daytime or nighttime conditions. The Canadian Ice Service (CIS) actively performs ice mapping generation and interpretation daily. Skilled ice analysts process SAR images to generate ice charts that have defined geographical regions, known as "polygons", with an assigned "egg code" to each polygon defined by the World Meteorological Organisation (WMO) [1]. The egg code contains numerical codes that define ice concentration by stage of development and floe size [9].

These ice charts have been used for decades but have limitations. First, egg codes are defined for large regions, so ice types are not identified at pixel resolution [3]. Second, polygons are generated subjectively. The results are affected by human bias, and there would be some variation between analysts. Third, the number of SAR scenes to be processed is expected to increase with time. CIS analysts used to process around ten scenes daily back to 2002 [4]. Nowadays, with the data provided by RADARSAT Constellation Mission (RCM) [5] and Sentinel-1, CIS interprets 110 (mid-March) to 290 (early September) SAR images every day. Manual interpretation of SAR imagery becomes very challenging due to the higher throughput required. Hence, manual procedures are potentially insufficient for future needs.

Therefore, an automated computer-based sea ice mapping method is desirable. A suitable operational automated method should have the following characteristics.

- The ability to generate ice maps that are consistent with expert interpretation.
- The ability to classify different ice types as required for operational use on full-scene SAR imagery.
- The ability to segment boundaries that match the natural ice and water boundaries.
- The ability to be invariant to SAR sensor artifacts such as banding noise and incidence angle variation.

However, mapping sea ice in SAR imagery is very challenging. First, different types of ice show very similar appearances in SAR imagery, especially when they are in contiguous development stages, such as young ice and first-year ice. Second, the ground truth is minimal. Supervised machine learning models rely on accurate pixel-level labels to achieve fair classification accuracy. However, ice charts released by primary national ice services only provide coarse labels for regions rather than pixels. Therefore, some studies prefer using samples selected from high-confidence regions for training and testing. Moreover, speckle noise and incidence angle variation can generate poor classification results. Boundaries between water and different types of ice, which are essential in sea ice maps, can be degraded. Although many studies [10, 11] report high numerical classification accuracy, The classification maps lack natural boundary information and suffer from speckle noise and inter-scan banding effect [12].

To address these challenges, we have developed and tested an automatic sea ice mapping method. The following summarizes the main contributions of this research.

1) We propose a novel sea ice classification method by integrating segmentation with pixel-based labeling. The method not only adopts texture features for classification but also preserves critical boundaries between water and different ice types. Unlike most existing methods that only focus on improving numerical classification accuracy, the proposed methods also aim to enhance the quality of classification maps.

2) To determine which classifier is more suitable for ice mapping, the performance of using support vector machine (SVM) and random forest (RF) in sea ice classification is compared. The results demonstrate that RF achieves better overall accuracy compared to SVM. Since previous research has not sufficiently compared the performance of the popular RF and SVM classifiers for sea ice classification, this benchmark would benefit other researchers in choosing the suitable classifier for their sea ice monitoring tasks.

3) To evaluate the robustness of the proposed method, we compared its performances using SAR images with and without applying incidence angle correction. The experiment results indicate that the proposed IRGS-RF is more robust to incidence angle variance than pixel-wise RF.

4) The methods are validated on a full-scene dataset covering the whole year, including melting, summer, and freezing seasons. The samples for validation are randomly selected across the full scene without any preferred regions. The experimental results demonstrate that the proposed method achieves accurate classification with well-preserved boundaries.

To the best of our knowledge, this is the first work that combines unsupervised segmentation with supervised labeling for open water as well as multiple ice types and validates it on a dataset that covers a whole year period. The rest of the paper is structured as follows. Section 2.2 provides a review of studies being done in sea ice classification. The dataset used in this paper is introduced in Section 2.3. Section 2.4 illuminates the steps of the proposed method. Section 2.5 presents the experimental results and analysis. Section 2.6 is the conclusion and future work.

2.2 Background

Significant research has been published in the last decades to explore automated sea-ice mapping systems based on SAR data. Early studies focused on modeling statistical distribution for ice types and water using backscattering intensity. Scheuchl et al. [13] explored the potential of using cross-polarization SAR imagery to monitor sea ice. The higher information content from dual-polarization data showed the capability to develop an automated sea ice classification system. Ward et al. [14] modeled the characteristics of ice and water using a mixture distribution. However, several studies concluded that only using backscatter intensity is insufficient in distinguishing different ice types [15, 16]. Therefore, many researchers turned to polarimetric SAR data since it holds more information separating different ice categories. Gill and Yackel [17] exploited the polarimetric parameters derived by decomposition algorithms with the maximum likelihood classifier to categorize different types of first-year ice. By extracting matrix invariant-based features from fully polarimetric ALOS-2 (L-band), Radarsat-2 (C-band), and TerraSAR-X (X-band) data, Singha et al. [18] separated water from sea ice with 100% accuracy.

Research has shown the potential of using quad-polarization SAR data for successful scene classification [17, 19, 18]. Nevertheless, the quad-polarization scene is not used operationally because of its narrow swaths. In contrast, dual-polarization data has been demonstrated to be a reliable source for sea ice-water classification when combined with textural features and machine learning methods. Many features have been explored for sea ice classification, e.g., Shannon entropy [20], local binary patterns [21], and cross-correlation between different polarizations [22]. A popular method for texture feature extraction from SAR sea ice images is the gray-level co-occurrence matrix (GLCM) [23]. Clausi [24] analyzed the relation between grey-level quantization and classification accuracy using the GLCM features. The study suggested that using contrast, entropy, and correlation with a quantization level of 64 is sufficient for classifying sea ice. Liu et al. [25] extracted GLCM features for segmentation and implemented an SVM to discriminate ice from water. Su et al. [26] combined surface temperature and GLCM features from the Moderate Resolution Imaging Spectroradiometer (MODIS) images to train an SVM model for ice-water classification. Tan [27] proposed a semi-automated ice mapping method and obtained a good identification for water.

Besides investigating different features, several studies have explored sea-ice classification using SAR imagery obtained at different working frequencies. Mahmud et al. [28] collected SAR data acquired from ALOS PALSAR (L-band), RADARSAT-2 (C-band), and QuikSCAT (Ku-band) to classify landfast first-year and multi-year ice in the Arctic. The results indicated that the L-band performed better for first-year ice, whereas the C-band is robust enough to distinguish multi-year ice. Given the longer wavelength, the L-band can detect the ice underneath melting ponds and wet snow in the melting season because of the enhanced penetration capability [29]. C-band is a better choice to monitor ice in the cold and dry winter since it provides details of surface roughness with higher resolution [30, 31].

To distinguish ice types, many classification models have been used, including the Bayesian classifier [32, 33], SVM [34], decision trees [35], and random forest [36, 37]. With the rapid development of graphics processing units (GPU) in the past decade, deep learning has been applied to remote sensing [38]. Ressel et al. [39] extracted GLCM-based textural features from TerraSAR-X imagery and fed them to a neural network, and classified three different ice types. Song et al. [40] combined a residual convolutional neural network (CNN) with long short-term memory (LSTM) units to learn spatial and temporal features of sea ice. Khaleghian et al. [10] compared the performance of several popular deep learning architectures for sea ice classification using Sentinel-1 data.

The studies mentioned above did achieve reasonable results. However, none of them has been deployed for operational sea ice classification for the following reasons. First, deep learning models usually perform inference on image patches rather than whole images. The patch size constrains the receptive field. A small patch size might provide insufficient characteristics for classification, while a large window size might contain different ice types and contaminate the information extracted. Thus, classification maps produced by pixel-wise deep learning models are usually contaminated by noise [41, 42]. Moreover, boundaries between ice and water can be smudged because of the inhomogeneity of the patch [43]. Similarly, GLCM features, since they depend on fixed-sized windows, also generate segmentation errors at class boundaries [44]. The defective classification result caused by this drawback is demonstrated and discussed in Section 2.5. Second, supervised machine learning methods require many reliable pixel-level labeled samples for training and testing. Since leading national agencies, such as CIS, the Norwegian Ice Service (NIS), and the Russian Arctic and Antarctic Research Institute (AARI), do not provide pixel-level ice charts, applying deep learning methods for operational use is not feasible at this time. When applying traditional machine learning methods with limited labeled samples has led to only training and validating on particular regions or sample points rather than full scenes [45, 46, 47].

As only using textural features with pixel-wise classifiers usually leads to poor-labeled sea ice classification maps [48], some researchers tried to refine the pixel-wise results to more visually appealing ice maps. Ochilov et al. [49] built a Markov random field (MRF) with maximum a priori (MAP) estimation for ice-water classification. Zhu et al. [45] first classified sea ice into five categories based on SVM. Then, a conditional random field (CRF) was applied to the original result as post-processing. Leigh et al. [50] combined textural and contextual features by modeling a CRF using a pixel-based classifier for ice-water classification. Collectively, these studies affirm that the integration of contextual information substantively improves the performance of sea ice classification.

Inspired by the research outlined above and the process employed by CIS to produce ice charts, this study seeks to establish a methodology that melds textural with contextual information for the discrimination of various sea ice types. Specifically, we will investigate which textural and contextual features should be learned and how these can be effectively integrated.

One of the challenges in sea ice mapping is addressing the effects of the SAR incidence angle [51]. Several studies [52, 53] have explored the potential of utilizing the SAR incidence angle as a feature to enhance classification accuracy. In this work, our primary goal is not to tackle the incidence angle effect directly through normalization or correction techniques. Nonetheless, it is worth noting that the region-based segmentation approach employed here proves effective in alleviating its adverse impacts [50]. Additionally, the texture-based features utilized in our study demonstrate resilience against variations in the incidence angle [54]. A more comprehensive discussion is provided in Section 2.4.

2.3 Dataset Used in the Study

2.3.1 Overview

This study is dedicated to developing sea-ice classification methods for mapping sea ice on pixel level in the Arctic region, currently under the purview of CIS's daily monitoring activities. Consequently, the selection of the dataset becomes a critical aspect in evaluating the performance of the proposed method and its comparison to other methods published in the literature.

With the burgeoning interest in sea ice monitoring, numerous research groups have taken the initiative to construct labeled sea ice datasets, subsequently making them publicly accessible to foster community-driven progress. Song et al. [55] constructed SI-STSAR-7, a labeled sea ice dataset encompassing 80 dual-polarization scenes captured by Sentinel-1. The authors proceeded to validate the reliability and effectiveness of this dataset by employing various models such as ConvLSTM [56], CNN [57], and SVM [58]. Wang and Li [59] introduced a U-Net based methodology to distinguish sea ice from open water, training and evaluating their model on a dataset comprising over 8,000 patches extracted from 251 Sentinel-1 dual-polarization scenes. In 2020, a collaborative endeavor between the Technical University of Denmark (DTU), the Danish Meteorological Institute (DMI), and the Nansen Environmental and Remote Sensing Center (NERSC) resulted in the release of the AI4Arctic/ASIP (ASID) sea ice dataset [60]. This dataset incorporates 461 Sentinel-1 HH and HV scenes spanning across the Greenland coast, accompanied by the respective ice charts. In late 2022, The Norwegian Computing Center, DMI, DTU, Polar View, NERSC, and the European Space Agency (ESA) jointly hosted a competition for sea ice mapping using a variant of the AI4Arctic dataset, called AI4Arctic sea ice challenge dataset [61]. The competition witnessed participation from over 100 teams, and the team from the University of Waterloo won the first prize.

Despite the datasets mentioned above being utilized in different research groups, they present certain limitations that pose challenges to their applicability in this study. First, this study aims to discriminate different sea ice types, not just distinguish open water from sea ice. Nevertheless, certain datasets [59] were specifically crafted for sea ice extent analysis, providing binary labels that offer only ice-water level information, thus proving inadequate for classifying different sea ice types. Second, the majority of the available datasets hinge on sea ice charts to generate labeled data for both training and testing purposes. However, ice charts typically only provide sea ice information at the polygon level, not at the pixel level. Consequently, these datasets employ varying sampling strategies to generate pixel-level labels from ice charts. For instance, SI-STSAR-7 solely considers polygons with a total ice concentration of 90% or more, where the dominant ice type also constitutes 90% or more. All pixels within a qualified polygon are assigned the same class. Given the dataset's predilection for samples from polygons with high confidence–an ice concentration exceeding 90%, it is unsuitable for evaluating the reliability and robustness of sea ice classification methods. Moreover, SI-STSAR-7 lacks scenes from melting seasons, which present a stark contrast to freeze-up seasons and pose additional challenges for operational tasks.

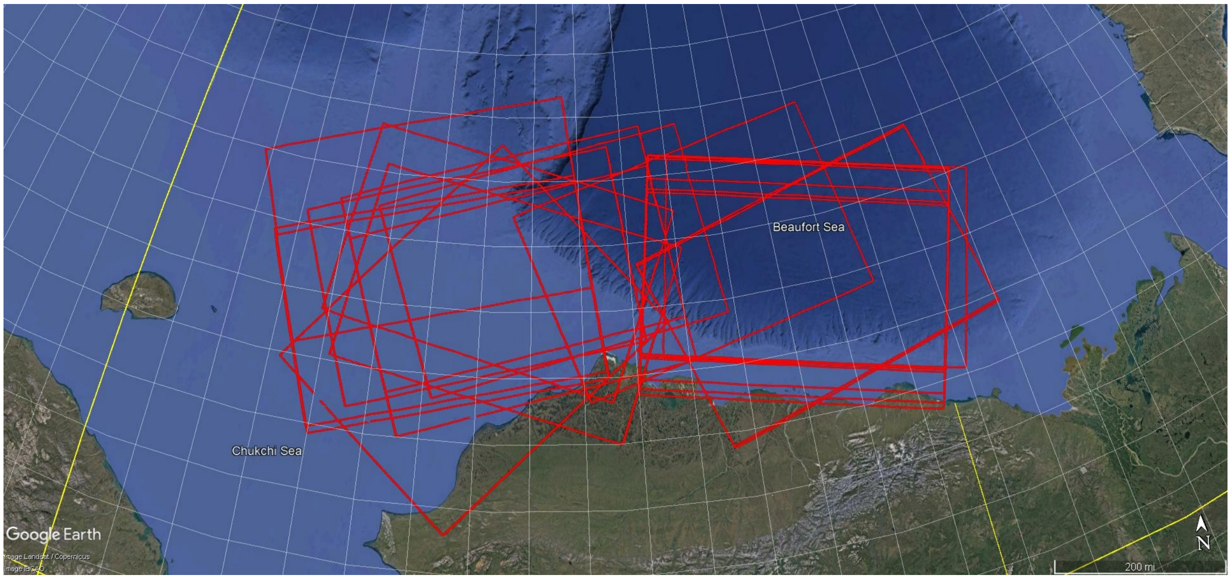


Figure 2.1: Location of the 18 SAR scenes used for evaluation in the Beaufort Sea.

Conversely, the AI4Arctic/ASIP scenes are distributed across a period spanning from March 14, 2018, to May 25, 2019, encompassing both freeze-up and melting seasons. However, ASID lacks explicit pixel-level labels. Relying on manually drawn ice charts, all pixels within a single polygon are assigned the same label, indicative of the dominant sea ice type. This absence of precise pixel-level labels hampers the generation of high-resolution sea ice maps using ASID.

2.3.2 Study Area and Satellite Data

Reflecting upon the discussion above, this study opts not to utilize any of the open-access datasets. Instead, we construct a new dataset, employing raw data previously utilized by Leigh et al. [50] for sea ice-water discrimination. This dataset, acquired in 2010, covers overlapping regions over the Beaufort Sea, falling under the monitoring jurisdiction of CIS. The geographical footprints of the dataset are depicted in Figure 2.1. The initial dataset comprises 20 scenes captured by the C-band RADARSAT-2 SAR satellite in a dual-polarized (HH and HV) ScanSAR wide beam mode. Due to the absence of detailed ice charts, a subset of 18 scenes is selected for the purpose of discriminating between sea ice types in this study. Table 2.1 enumerates the capture dates for each scene. The dataset boasts a nominal pixel spacing of 50 by 50 m, with image dimensions approximately measuring 11,000 by 10,000 pixels, representing the largest range size attainable by RADARSAT-2. Each scene covers 500 km in both the azimuth and range dimensions. The incidence angle spans from 20°to 50°, encompassing both ascending and descending orbits. The data acquisition period includes both melting and freeze-up seasons, identified as the most challenging periods of the year. It is important to note that sea ice conditions from January through March tend to be more stable and are of lesser importance for operational purposes, in addition to posing fewer challenges for scene classification. Hence, the dataset deliberately excludes this period.

Table 2.1: Dataset used in this study. The date of capture, satellite orbit: descending (D)/ascending (A), and incidence angle (IA) range are included .

Scene ID	Date of capture	Orbit	IA (°)
20100418	20100418 16:33:15	D	19.72 to 49.46
20100426	20100426 04:04:39	A	19.55 to 49.44
20100510	20100510 03:56:20	A	19.58 to 49.44
20100524	20100524 03:47:56	A	19.61 to 49.46
20100623	20100623 04:12:55	A	19.63 to 49.45
20100629	20100629 16:33:26	D	19.71 to 49.43
20100721	20100721 17:32:08	D	19.64 to 49.47
20100730	20100730 16:29:08	D	19.61 to 49.46
20100807	20100807 17:36:10	D	19.77 to 49.47
20100816	20100816 16:33:29	D	19.74 to 49.45
20100907	20100907 03:56:14	A	19.59 to 49.44
20100909	20100909 16:33:21	D	19.63 to 49.48
20101003	20101003 16:33:24	D	19.59 to 49.46
20101021	20101021 04:13:25	A	19.50 to 49.43
20101027	20101027 02:57:26	A	19.58 to 49.43
20101114	20101114 04:13:04	A	19.57 to 49.43
20101206	20101206 01:51:39	A	19.59 to 49.39
20101214	20101214 02:57:25	A	19.58 to 49.45

Figure 2.2 shows an example scene from the dataset, captured on April 26, 2010. Figures 2.2(a) and 2.2(b) display the HH and HV scenes, respectively, with histogram stretching applied for visual enhancement. This particular scene is characterized by its complexity, showcasing different types of ice, including both first-year and multi-year ice.

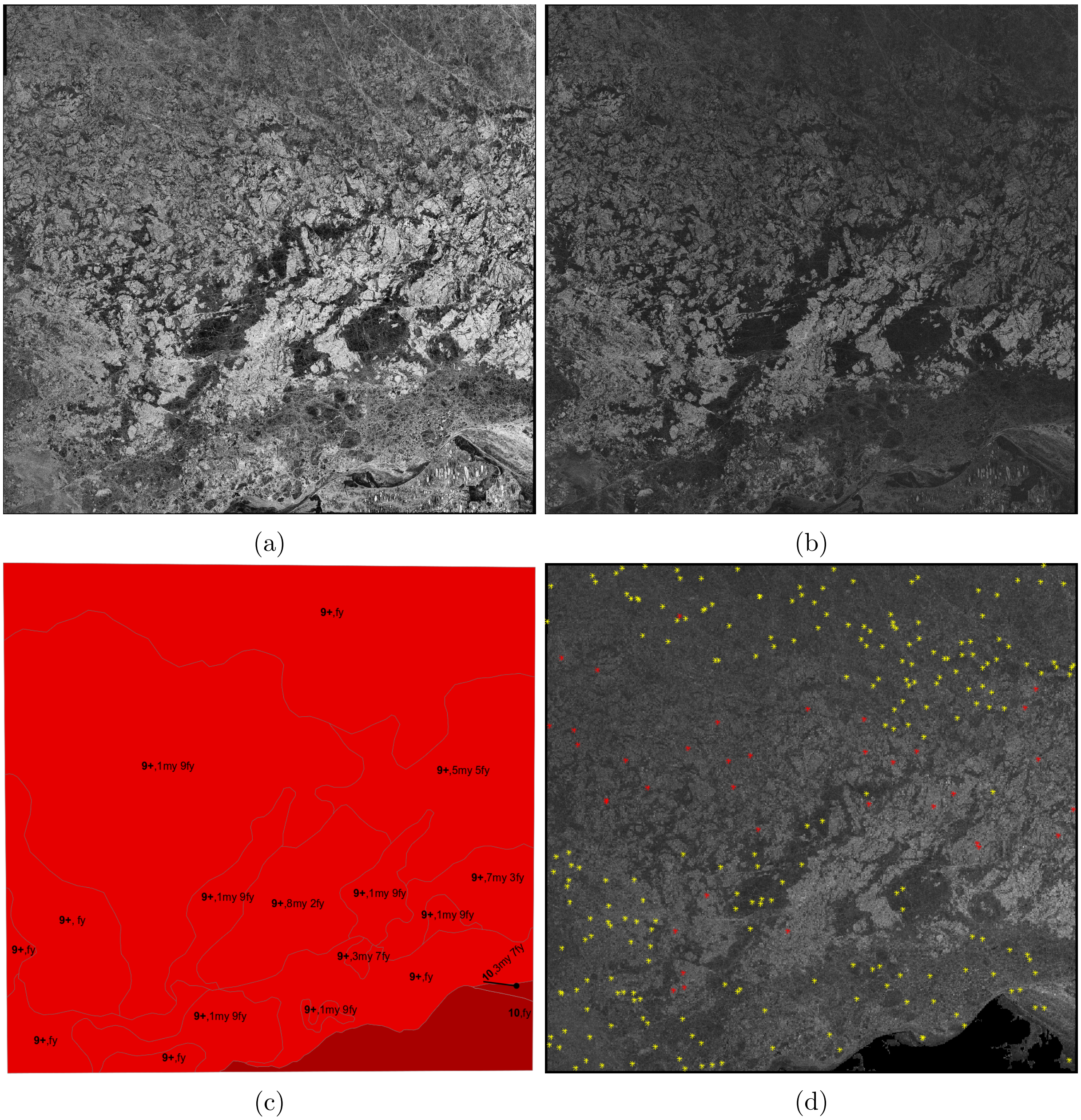


Figure 2.2: Example scene captured on April 26, 2010 (Scene ID: 20100426_040439). (a) HH polarization. (b) HV polarization. (c) Reference ice chart. (d) 500 sample points selected for training and testing (Yellow: first-year ice. Red: multi-year ice).

2.3.3 Sea Ice Charts

The ice chart, released by CIS in Canada, stands as the most authoritative data source for sea ice information. This chart is crafted by analysts through a two-step process. Initially, regions characterized by a similar concentration and a predominant ice type are grouped together to form a polygon. Following this, the polygon is annotated with a metric that encapsulates both the overall ice concentration and the specific sea ice conditions.

CIS releases sea ice charts on a daily or weekly basis. Therefore, temporal discrepancies may arise between the sea-ice conditions depicted in the scene and the corresponding ice charts. Furthermore, the ice charts themselves are of coarse resolution. To acquire more accurate and detailed ice charts that align with the SAR scenes in our dataset, a former sea ice analyst from CIS volunteered his expertise, contributing to the production of the sea ice charts utilized in this study.

In this study, we focus on classifying four distinct sea ice types (stage of development), as defined by WMO [62], open water (OW), young ice (YI), first-year ice (FYI), and multi-year ice (MYI). Figure 2.2(c) illustrates the sea ice chart corresponding to the SAR scenes depicted in Figures 2.2(a) and 2.2(b). The annotations on each polygon provide details on the total ice concentration, the various types of sea ice present, and their respective proportions of the total concentration. Depending on the diversity of sea ice types within a polygon, the annotation could comprise two or three components, separated by commas. The first component is a digit ranging from 0 to 9, increasing in increments of 1, representing the total ice concentration. Here, '0' symbolizes 0% concentration (open water), and '9' stands for 90% concentration, while '9+' denotes concentrations exceeding 90% (virtually no open water). The subsequent components, where the third one is optional, specify the types of sea ice and their proportions in the total concentration. The

For instance, an annotation such as '9+, 1my 9fy' on a central polygon indicates a sea ice concentration surpassing 90%, with a composition of 10% multi-year ice and 90% first-year ice. Another example, '9+, fy', conveys a concentration above 90%, exclusively comprising first-year ice.

2.3.4 Pre-processing

Upon acquisition from MDA Ltd, the original SAR data is georeferenced and stored in GeoTIFF format, encapsulating the digital numbers represented in intensity for image pixels. The dataset also includes output scaling lookup tables (LUTs) for sigma-nought, beta-nought, and gamma. The digital numbers in SAR images are calibrated to sigma-nought using the LUT in this study.

The original SAR image sizes in the dataset are originally of a substantial size, approximately 10,000 $\times$ 10,000 pixels. Previous studies have demonstrated that classification results derived from downsized images still meet the requirements for the operational sea ice maps, as long as the downsampling window size remains relatively small, specifically less than 10 $\times$ 10 [59, 63, 64]. In light of this, we chose to downsample our SAR images using a 4 $\times$ 4 averaging window. While we have not applied any filters specifically to address speckle noise, this downsampling process inherently aids in mitigating its impact on the SAR images.

2.3.5 Training and Testing Samples

The development of machine-learning-based classification methods for SAR images crucially depends on the availability of ample training data, accurately labeled down to the pixel level. In this study, where the goal is to create pixel-level sea ice maps, it becomes crucial that both the training and testing samples possess precise labels. To generate a dataset that meets these criteria, we randomly selected 500 sample pixels across the entirety of each scene, deliberately avoiding land areas and ensuring no bias towards any particular region. We maintained the assumption that each pixel in our dataset represents a single type of sea ice, leading us to exclude pixels located on the boundaries between different sea ice types from our selection process. The corresponding ice chart served as the reference for labeling these sample pixels. Figure 2.2(d) provides a visual representation of a scene annotated with labeled sample points. Based on the ice chart, this particular scene contains both first-year and multi-year ice, with the 500 labeled samples highlighted in yellow (FYI) and red (MYI).

To conduct an exhaustive assessment of both the performance and the robustness of the proposed method, we employ the leave-one-out (LOO) cross-validation strategy for the training and testing phases. In this approach, each test scene is evaluated separately, with the training dataset comprising all 8,500 labeled samples drawn from the other 17 scenes in the collection. Subsequently, the method is tested on a set of 500 samples originating from the test scene itself. This evaluation strategy ensures a comprehensive understanding of the method's capabilities and its consistency across different scenes.

2.4 Methodology

Based on the extensive literature review presented in Section 2.2, it is evident that supervised classification methods predominantly focus on learning textural features, resulting in reasonably accurate pixel-level results. Nonetheless, these methods exhibit susceptibility to noise and variations in incidence angle. On the other hand, segmentation algorithms excel at extracting contextual information and efficiently dividing SAR images into homogeneous regions that align closely with natural boundaries. Despite their strengths, these segmentation algorithms fall short when it comes to accurately assigning labels to homogeneous regions.

Several studies have delved into the integration of textural and contextual features to enhance sea ice classification. However, there remains a substantial gap in understanding how to effectively extract these two types of information and, more crucially, the methodologies for their successful integration during the classification process. Addressing this gap, this study endeavors to establish a comprehensive infrastructure designed to synergistically combine both textural and contextual information, aiming to significantly improve the accuracy and reliability of sea ice classification in SAR images.

2.4.1 Problem Formulation

Let Y denote the SAR image that consists of N pixels, i.e., $Y = \{y_i | i = 1, 2, ..., N\}$ and L is the associated label map $L = \{l_i | i = 1, 2, ..., N\}$ that consists of a total of K classes of different ice types, where $l = \{1, 2, ..., K\}$. Sea ice mapping from SAR image aims to estimate L given Y.

Algorithmically, first, Y is segmented into a total of T homogeneous regions $Y = \{R_1, R_2, ..., R_T\}$, where R_r consists of n_r pixels:

$$R_r = \{y_i^r | i = 1, 2, ..., n_r\} \tag{2.1}$$

To estimate the label of R_r, denoted by l_{R_r}, the label of each pixel, i.e., $\{l_{y_i} | i = 1, 2, ..., n_r\}$, in R_r is estimated. Then, $\{l_{y_i}\}$ is used to derive the label of the region l_{R_r} via the proposed energy function.

The proposed classification system consists of two main components shown in Figure 2.3. The system uses HH and HV polarized images of the scene, a pre-trained pixel-wise classifier, and an optional landmask file as inputs. The left block in the flowchart

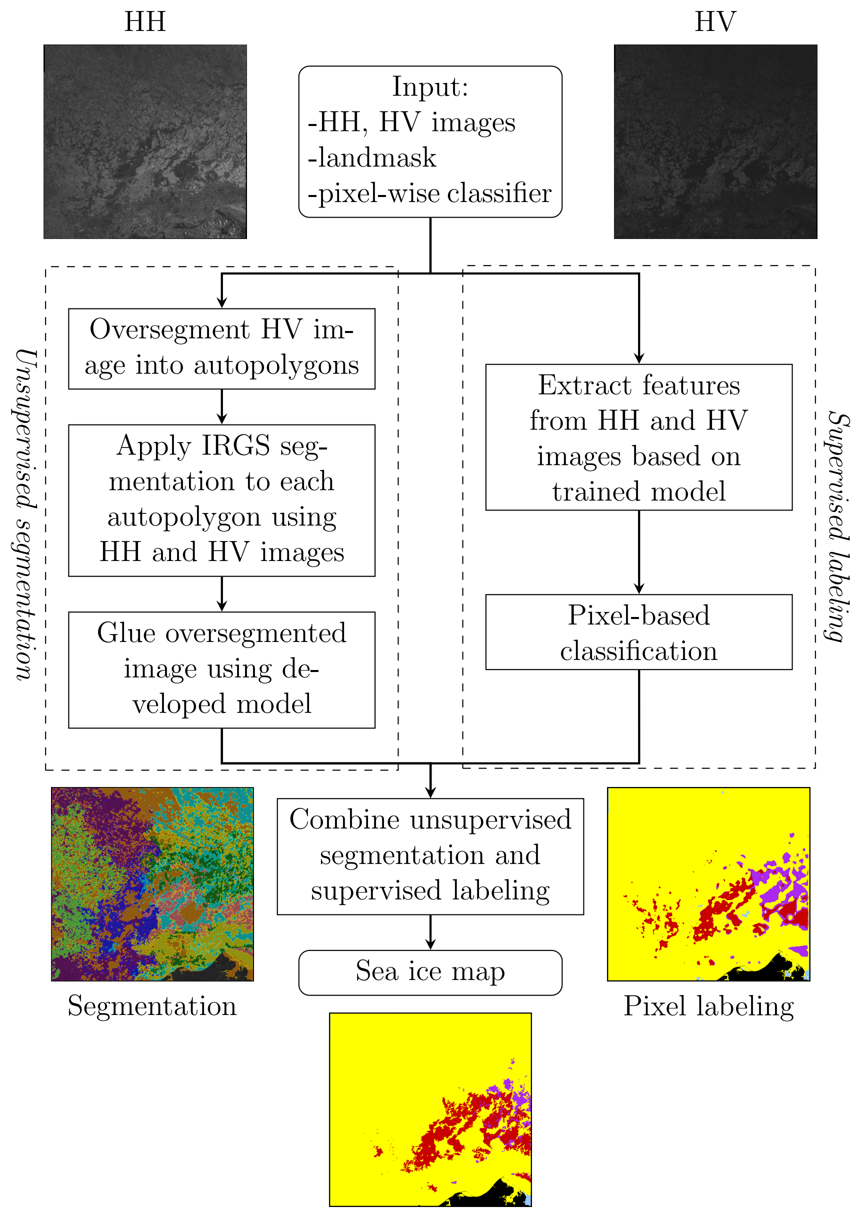


Figure 2.3: Flowchart of the proposed ice mapping system. Inputs are HH/HV images, landmask (optional), and a trained classifier (SVM or RF). The left block calculates contextual information by unsupervised segmentation, while the textural feature is extracted in the right block. Then, these features are combined using the proposed energy function to generate the final classification map.

is the Iterative Region Growing with Semantics (IRGS) segmentation [65] to generate $Y = \{R_1, R_2, ..., R_T\}$, and the right block is the pixel-wise labeling to determine l_{y_i}. Details are described in the following subsections.

2.4.2 Unsupervised Segmentation

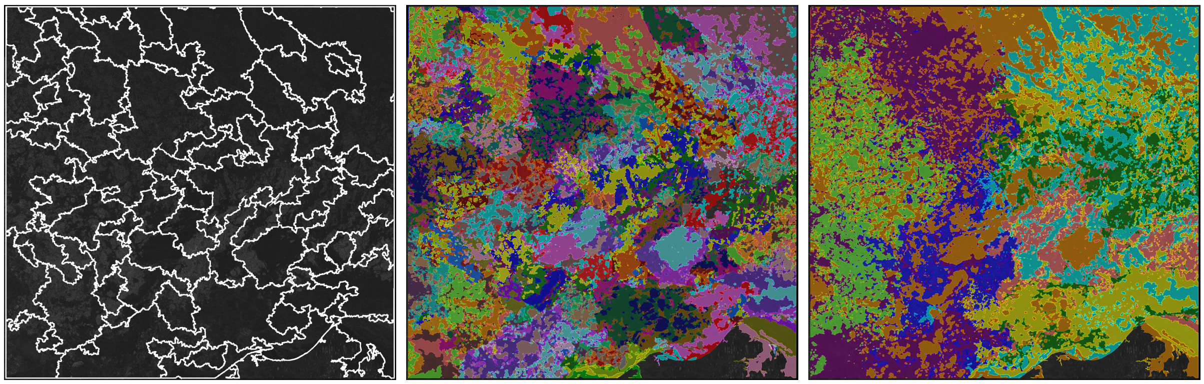
Numerous methodologies have been developed for image segmentation. Compared with alternative segmentation algorithms, such as SLIC [66], watershed [67], GraphCut [68], MRF [69], the IRGS algorithm is specifically tailored for SAR images. It incorporates edge strength with a spatial context model to segment images using an iterative region-growing strategy. Previous studies have affirmed that the segmentation results carried out by IRGS are robust to speckle noise and variations in incidence angle [65, 70, 71, 72, 73]. Consequently, IRGS has been selected to extract contextual information for sea ice analysis in this study.

Operational SAR imagery used for sea ice mapping at CIS has large extents. The gradual change of incidence angle from near range to far range leads to a corresponding change in within-class backscatter. Therefore, a two-step segmentation strategy called 'glocal', shown in Figure 2.3, was introduced to suppress the incidence angle effect [50]. First, the whole scene is segmented into sub-regions called 'autopolygons' [74] using a modified watershed algorithm [75]. Only the HV scene is used in this step, which is shown in Figure 2.4(a), because it is less sensitive to both incidence angle variation and surface roughness caused by winds [76].

Within each autopolygon, an IRGS segmentation is performed using the HH and HV polarized images. This step is presented in 2.4(b). Each region in an autopolygon results from oversegmention and is regarded as a node in a region adjacency graph (RAG). Since each node is homogeneous and only contains water or one ice type, an arbitrary label is assigned to each region for further processing. This is the local step of the unsupervised segmentation. The effects of speckle noise and incidence angle are restrained by processing each autopolygon individually.

The second step is called 'global'. Once the oversegmented result is available, a gluing step is operated across the whole image. The edge strength and statistical information from HH and HV polarization are considered during this merging step. The final segmentation result contains six classes with arbitrary labels. This method is called 'glocal' as it combines local oversegmention and global merging.

The structure of IRGS segmentation is elucidated in [70]. Each autopolygon is segmented into four clusters, and the clusters are merged into six classes in the global step.



(a) (b) (c)

Figure 2.4: IRGS segmentation results for April 26, 2010 (scene ID 20100426). (a) Autopolygons generated by the modified watershed algorithm. (b) Local oversegmentation result. (c) Glocal results.

The following parameters are applied to the segmentation for all the scenes in the dataset. β_1 and β_2 used for estimating multilevel logistic model (MLL) are 3 and 0.4, respectively. The number of iterations is set to 100 to achieve an over-segmented result. According to previous experiment results, the values of β_1 and β_2 have little impact on the final over-segmentation results when the iteration reaches 100.

2.4.3 Supervised labeling

Features

The GLCM features serve as statistical tools for analyzing an image's texture, proving to be highly effective in various applications, including SAR image classification and, more specifically, sea ice classification. Sea ice, with its rich texture information, makes GLCM features particularly relevant for these applications. In this study, we harness the potential of GLCM features to construct an automated sea ice classification system. The specific GLCM features [77] used are defined as follows:

- ASM: Angular second moment

$$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} P_{ij}^2 \tag{2.2}$$

- CON: contrast

$$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} P_{ij}\left(i-j\right)^2 \tag{2.3}$$

- DIS: dissimilarity

$$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} P_{ij}\left|i-j\right| \tag{2.4}$$

- ENT: entropy

$$-\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} P_{ij}\log P_{ij} \tag{2.5}$$

- HOM: homogeneity

$$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} \frac{P_{ij}}{1+\left(i-j\right)^2} \tag{2.6}$$

- INV: inverse moment

$$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} \frac{P_{ij}}{1+\left(i-j\right)} \tag{2.7}$$

- MU: mean

$$\mu = \sum_{i=0}^{N-1} i \sum_{j=0}^{N-1} P_{ij} \tag{2.8}$$

- STD: standard deviation

$$\sigma = \sqrt{\sum_{i=0}^{N-1}\left(i-\mu\right)^2 \sum_{j=0}^{N-1} P_{ij}} \tag{2.9}$$

- COR: correlation

$$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} \frac{\left(i-\mu\right)\left(j-\mu\right)P_{ij}}{\sigma^2} \tag{2.10}$$

where P_{ij} represents the probability of co-occurrence of the gray levels i and j in a N by N GLCM. N denotes the number of distinct gray levels in the image, while μ and σ^2 represent the mean and the variance of the GLCM, respectively. The angle is fixed at 0°

The choice of window and step size of GLCM features can impact the performance of sea ice classification. The window size determines the perceptive area for textural feature extraction. For instance, a small window usually works for distinguishing open water since the calm water surface has less texture in contrast to different types of sea ice. Small window sizes also work better to detect textural features from within leads and floes, while the complex repeating patents caused by fissures and cracks in different ice types require larger window sizes to capture. The spatial distance of GLCM features determines the scale of repeating patterns. For example, first-year ice has more dense repeating patterns compared with multi-year ice. The chosen window and step sizes of GLCM features are listed in Table 2.2. In addition to the 162 GLCM features, we add individual pixel intensity, local average, and maximum pixel intensities in 5 X 5 and 25 X 25 windows. All the features are extracted from HH and HV polarized scenes, resulting in a set of 172 features [50].

Table 2.2: GLCM parameters used in the study.

Window size (pixels)	**Spatial distance (pixels)**
5 by 5	1
11 by 11	1
25 by 25	1
25 by 25	5
51 by 51	5
51 by 51	10
51 by 51	20
101 by 101	10
101 by 101	20

In order to minimize computation time and to minimize the 'curse of dimensionality' [78, 79], a feature search was performed to reduce the number of features. Recursive feature elimination with cross validation [80] is applied to select the best feature combination in this study. The feature with the least importance is discarded in each iteration. The process is repeated until the best feature combination is found. Since the dataset used in this work consists of 18 scenes, the feature search is deployed with a cross-validation strategy. The feature search executes 18 times. In each loop, a feature importance estimator is trained on the 17 scenes and tested on the remaining scenes to determine the importance of each feature. This procedure was carried out 18 times, and each scene was enrolled in both training and test sets. This cross-validation strategy is called leave-one-out (LOO). After the LOO is performed, the feature importances from each iteration are summed up to calculate the final feature ranking. After running the feature search, the 30 most important

Table 2.3: List of selected 30 features using RFE and cross-validated scheme.

#	Polarization	Feature Type	Window size	Step size
1	HH	Pixel Intensity	N/A	N/A
2	HV	Pixel Intensity	N/A	N/A
3	HV	GLCM MU	11 by 11	1
4	HH	GLCM COR	25 by 25	1
5	HH	GLCM MU	25 by 25	1
6	HV	GLCM MU	25 by 25	1
7	HH	GLCM MU	25 by 25	5
8	HV	GLCM MU	25 by 25	5
9	HH	GLCM MU	51 by 51	5
10	HV	GLCM ASM	51 by 51	5
11	HV	GLCM HOM	51 by 51	5
12	HV	GLCM MU	51 by 51	5
13	HH	GLCM MU	51 by 51	10
14	HV	GLCM ASM	51 by 51	10
15	HV	GLCM MU	51 by 51	10
16	HH	GLCM MU	51 by 51	20
17	HV	GLCM ASM	51 by 51	20
18	HH	GLCM DIS	101 by 101	10
19	HH	GLCM INV	101 by 101	10
20	HH	GLCM MU	101 by 101	10
21	HV	GLCM ASM	101 by 101	10
22	HV	GLCM HOM	101 by 101	10
23	HV	GLCM INV	101 by 101	10
24	HV	GLCM MU	101 by 101	10
25	HH	GLCM CON	101 by 101	20
26	HH	GLCM MU	101 by 101	20
27	HV	GLCM ASM	101 by 101	20
28	HV	GLCM HOM	101 by 101	20
29	HV	GLCM INV	101 by 101	20
30	HV	Pixel Average	25 by 25	N/A

features are selected. The result is listed in Table 2.3.

Selecting an appropriate classifier is paramount for the success of a classification task.

In their comprehensive comparison of various classification techniques using RADARSAT-1 imagery, Xu et al. [81] demonstrated that tree-based classifiers and SVM outperformed others, such as artificial neural networks, generalized additive models, and penalized linear discriminant analysis. SVM, in particular, is renowned for its proficiency in managing high-dimensional data, which is a common characteristic of SAR images. This trait becomes even more pertinent when incorporating GLCM features, as is the case in this study.

Simultaneously, RF emerges as another compelling choice due to its versatility in handling high-dimensional spaces and its robustness against overfitting. This robustness is particularly beneficial given the LOO cross-validation strategy employed in our evaluation process. Consequently, both RF and SVM have been selected as the classification methods for this study, with the anticipation that they will contribute significantly to the effectiveness and accuracy of the sea ice classification task.

Support vector machine (SVM)

A support vector machine is a supervised learning model employed for classification and regression. In SVM, the objective is to find a hyperplane in a high-dimensional space that distinctly separates different classes. However, the decision boundary is determined by only a subset of the training samples, known as support vectors, which are the data points closest to the hyperplane and crucial for defining the margin. These support vectors are essential as they are the most challenging to classify, providing a robust and generalized decision boundary. The training process of SVM is to minimize the following loss function.

$$\left[\frac{1}{n}\sum_{i=1}^{n}\max\left(0,1-y_i\left(\mathbf{w}^T\mathbf{x}_i-b\right)\right)\right]+\lambda\left\|\mathbf{w}\right\|^2 \tag{2.11}$$

where $\mathbf{w}$ is the weight vector, b is the bias, $y_i \in -1, 1$ are the class labels, and λ is a regularization parameter. Note that SVM is inherently a binary classifier. Therefore, for multi-class problems like distinguishing different sea ice types, a one-vs-rest or one-vs-one strategy would be necessary. This study applied the one-vs-rest strategy with label $y_i \in \{1, 2, 3, 4\}$ assigned to YI, FYI, MYI, and OW, respectively.

The standard SVM assumes linear separability in the data. However, many real-world problems present non-linear patterns. To tackle this, kernel SVMs use kernel functions, such as polynomial, radial basis function (RBF), and sigmoid, to implicitly map the input data to a higher-dimensional space where a linear separator might exist. In this work, we utilized RBF kernel defined as:

$$K\left(\overline{t_i}, \bar{t}\right) = \exp\left(-\gamma \left|\overline{t_i} - \bar{t}\right|^2\right) \tag{2.12}$$

where γ is a parameter that determines the scale of the Gaussian function, and $\overline{t_i}$ and $\bar{t}$ are input samples. The RBF kernel effectively transforms the data into a higher-dimensional space, making it possible to find a linear separator. Due to the sparse nature of SVMs, where the decision boundary is defined by a subset of the training samples (support vectors), they require less memory and computational power, mitigating the risk of overfitting. Furthermore, SVMs are versatile and adaptable to various classification tasks with different kernel functions [82].

Random Forest

RF [83] is an ensemble learning model for classification and regression. It is an aggregation of diversified decision trees bound by a bootstrap aggregating (bagging) strategy. The key idea of RF is random–both samples and features are selected to train the decision trees and their nodes. Each decision tree is trained by a subset of the whole dataset using bootstrap sampling, and the node of the tree is grown from a random feature of the input data. This learning method averages the learned weights for all the features and helps to prevent overfitting problems. The bagging method makes the classifier tolerate the noise appearing in the dataset, which is crucial for our research since SAR images are contaminated by speckle noise. The aggregation of decision trees can be trained by parallel computing and reduce the training and testing time [84]. In contrast to SVM, RF does not require feature searching since features are randomly selected, and the importance is assigned to each feature using GINI index [85]. The hyper-parameters of RF are determined by utilizing cross-validation-based grid searching. The search range for the number of trees is $\in \{25 \to 500 \mid \text{step} = 10\}$, max depth is $\in \{2 \to 20 \mid \text{step} = 2\}$, and minimum samples per leaf is $\in \{1 \to 10 \mid \text{step} = 1\}$. The best combination of hyper-parameters determined by grid search is: the number of trees=200, max depth=12, and minimum samples per leaf=2. In order to evaluate the performance of SVM and RF using the same feature set, RF is trained on the selected features in this paper.

In this study, SVM and RF are deployed as pixel-wise classifiers for the following reasons. First, both models have been applied for remote sensing tasks, and the results are promising, according to previous studies. Second, SVM and RF have their own techniques to suppress overfitting and increase robustness. Finally, the runtime for predicting each scene is less than 30 minutes, which is acceptable for operational use.

2.4.4 Combination of segmentation and labeling

By performing the unsupervised segmentation algorithm, $Y = \{R_1, R_2, ..., R_T\}$ is determined. Each R_r in Y is homogeneous and is assigned an arbitrary label. For the results carried out by pixel-wise classifiers, SVM and RF, each pixel is labeled as an ice type or water. The flow chart of the proposed automatic sea ice mapping system is illustrated in Figure 2.3. The system ingests the HH and HV polarized images of the scene, a pre-trained pixel-wise classifier, and an optional landmask file to neglect the land and image boundary. The pixel-wise classifier, which is SVM or RF in this paper, is trained using the selected 30 features. Inspired by the mechanism of RF, an energy function $E(k)$ is proposed to obtain l_{R_r} using $\{l_{y_i} | i = 1, 2, ..., n_r\}$.

$$E(k) = \frac{1}{n_r} \sum_{i=1}^{n_r} \sum_{k=1}^{K} w(k, l_{y_i}) \tag{2.13}$$

$$w(k, l_{y_i}) = \begin{cases} 0 & l_{y_i} = k \\ 1 & l_{y_i} \neq k \end{cases} \tag{2.14}$$

$$l_{R_r} = \text{arg min} \left(E(k) \right) \tag{2.15}$$

where n_r is the number of pixels in the homogeneous region R_r. w is the weight function and $k \in \{1, 2, ..., K\}$ is the class label. When combining the unsupervised segmentation with supervised labeling, all the pixels in the same region share the same label, which is determined as the dominant class according to the pixel-wise classification result, with all pixels within it.

2.5 Experiments and Analysis

As highlighted in Section 2.4, the proposed infrastructure is designed to be compatible with any pixel-wise classifier, making both SVM and RF suitable for integration. Consequently, two combined models–IRGS-SVM and IRGS-RF–have been implemented and tested to assess the performance and reliability of our proposed framework. Additionally, SVM and RF were selected as benchmark methods to provide a basis for comparison, helping to elucidate the enhancements facilitated by our infrastructure. This section presents and discusses both numerical and visual results obtained from these experiments.

Previous literature [55, 86, 64] has established certain expectations regarding the performance of pixel-wise classifiers like SVM and RF. Typically, these classifiers achieve satisfactory numerical accuracy; however, the resulting sea ice maps are marred by noise-like classification errors, primarily due to a lack of contextual information. In contrast, while the numerical classification accuracies of IRGS-SVM and IRGS-RF are anticipated to be slightly higher, substantial improvements are expected to manifest in the produced sea ice maps. These enhancements include a dramatic reduction in classification errors and preservation of natural boundaries between different sea ice types.

To evaluate the robustness and generalization capabilities of the proposed methods, a Leave-One-Out (LOO) cross-validation strategy has been employed. Under this scheme, each scene is classified using a model trained on the remaining 17 scenes in the dataset, with this process iterated 17 times to ensure complete separation of training and test samples. Based on a set of 500 reference pixels per scene, the overall accuracy results are tabulated in Table 2.4. The pixel-based SVM and RF classifiers yielded average overall accuracies of 81.13% and 83.84%, respectively. Both classifiers encountered difficulties when applied to scenes from the freeze-up season, particularly evident in the results from November 14, 2010. In this challenging scene, SVM and RF achieved accuracies of just 60.50% and 65.40%, respectively. Conversely, the least challenging scene, recorded on September 7, 2010, featured exclusively open water, significantly simplifying the classification task.

After integrating with the segmentation results based on the proposed framework, the IRGS-SVM and IRGS-RF achieved average overall accuracies of 84.07% and 86.33%, with improvements of 2.94% and 2.49% compared with pixel-wise SVM and RF, respectively. The confusion matrices of SVM, IRGS-SVM, RF, and IRGS-RF are shown in Table 2.5, 2.6, 2.7, and 2.8, respectively. All four classifiers distinguish water from ice with accuracies over 94%. Even with the interference caused by wind and waves, the features are sufficient to describe their characteristics. However, SVM and RF struggle to separate young ice from other classes. A possible reason for the misclassified young ice may lie in a similar backscattering to first-year ice. 34.45% of young ice is categorized as first-year ice by RF. Young ice only appears in five scenes in the dataset, covering from late October through December. The limited labeled samples may cause poor performance in classifying young ice. Although pixel-based classifiers do not achieve satisfactory results, the proposed IRGS-SVM and IRGS-RF are able to boost the classification accuracy and present visually appealing sea ice maps.

Several studies have suggested that applying corrections for incidence angle variations can lead to improvements in classification accuracy for synthetic aperture radar (SAR) imagery [29, 52]. This relationship is demonstrated by the previously established correlation between backscatter coefficients and incidence angle [31, 34]. In our study, we leverage

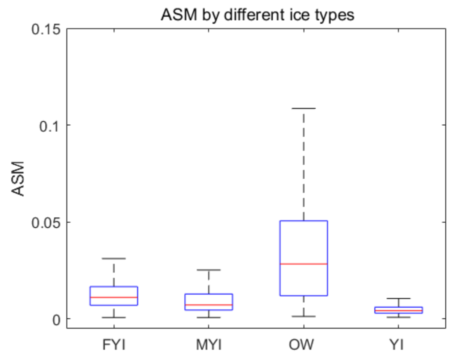
Table 2.4: Classification result for all 18 scenes.

Scene ID	SVM	RF	IRGS-SVM	IRGS-RF
20100418	80.30%	**87.60%**	83.20%	86.10%
20100426	83.40%	88.00%	86.40%	**88.10%**
20100510	85.40%	87.30%	85.90%	**88.00%**
20100524	75.60%	79.00%	85.10%	**86.30%**
20100623	88.10%	**90.80%**	90.20%	90.20%
20100629	70.20%	71.70%	82.40%	**84.10%**
20100721	87.40%	87.40%	85.60%	85.50%
20100730	**78.70%**	73.70%	75.20%	78.10%
20100807	91.90%	93.30%	94.70%	**95.30%**
20100816	86.20%	**86.80%**	85.00%	85.00%
20100907	98.50%	100.00%	100.00%	100.00%
20100909	98.20%	**98.80%**	98.00%	98.00%
20101003	77.70%	77.20%	77.80%	77.70%
20101021	89.40%	90.80%	92.60%	**95.60%**
20101027	65.10%	72.20%	73.00%	**76.90%**
20101114	60.50%	68.80%	65.40%	**75.60%**
20101206	74.90%	83.00%	81.40%	**86.60%**
20101214	69.80%	72.70%	71.80%	**76.90%**
Overall	81.13%	83.84%	84.07%	**86.33%**

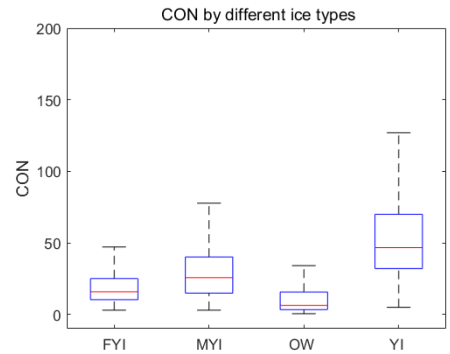
these insights to estimate incidence angle dependencies for each sea ice type. Given that the incidence angle in our dataset varies from 20°to 50°, we opted to normalize the sigma-nought values using a reference incidence angle of 35°.

To assess the impact of incidence angle correction on classification accuracy, we conducted experiments using RF and IRGS-RF, which demonstrated superior performance compared to their counterparts. These experiments were performed on HH/HV scenes to investigate whether the incidence angle correction contributes to better classification results. The comparison of classification accuracies, both before and after the application of incidence angle correction, is presented in Table 2.9.

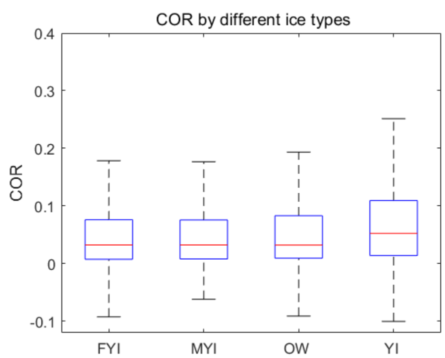
The results indicate a 0.66% improvement in overall accuracy for the RF classifier on the normalized dataset. However, it is crucial to note that a decline in accuracy was observed in several scenes. This discrepancy could potentially be attributed to the weakening of texture patterns for specific GLCM features in the near range, a phenomenon consistent



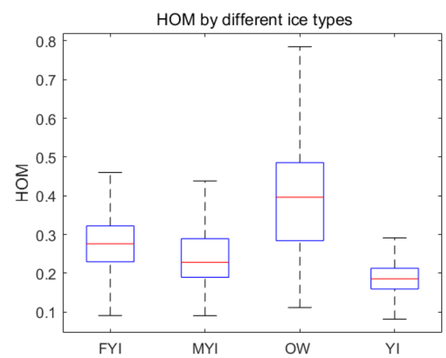
(a)



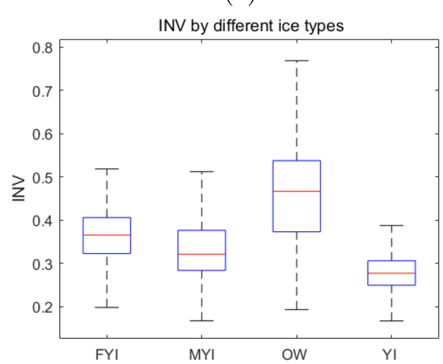
(b)



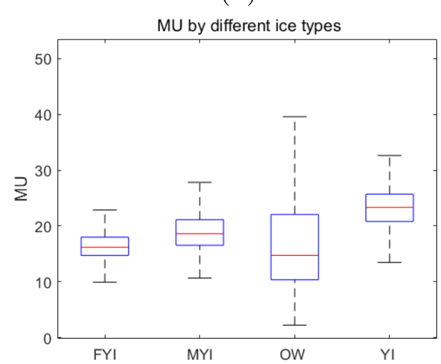
(c)



(d)



(e)



(f)

Figure 2.5: The distribution of GLCM features for different sea ice types and open water in box-and-whisker plots. (a) ASM. (b) CON. (c) COR. (d) HOM. (e) INV. (f) MU. All the GLCM features are calculated using a size of 51 with a stride of 20.

Table 2.5: Classification confusion matrix of SVM.

	YI	FYI	MYI	OW
YI	**44.86%**	40.50%	4.29%	10.35%
FYI	5.16%	**85.27%**	4.12%	5.46%
MYI	16.56%	5.46%	**70.25%**	7.73%
OW	0.50%	4.72%	0.02%	**94.76%**

Table 2.6: Classification confusion matrix of RF.

	YI	FYI	MYI	OW
YI	**57.52%**	34.45%	5.17%	2.86%
FYI	3.97%	**87.68%**	2.61%	5.74%
MYI	13.47%	13.28%	**72.07%**	1.18%
OW	1.18%	4.60%	0.06%	**94.16%**

Table 2.7: Classification confusion matrix of IRGS-SVM for all 18 scenes.

	YI	FYI	MYI	OW
YI	**58.747%**	29.17%	0.48%	11.62%
FYI	3.06%	**91.16%**	1.53%	4.25%
MYI	13.58%	2.46%	**77.30%**	6.65%
OW	0.49%	4.98%	0.00%	**94.53%**

Table 2.8: Classification confusion matrix of IRGS-RF for all 18 scenes.

	YI	FYI	MYI	OW
YI	**70.77%**	22.74%	3.21%	3.28%
FYI	2.56%	**91.92%**	1.79%	3.74%
MYI	8.20%	10.03%	**78.85%**	2.92%
OW	0.67%	4.84%	0.03%	**94.46%**

with findings reported in previous studies [54, 87, 88]. On the other hand, the IRGS-RF classifier demonstrated robustness to variations in incidence angle, with negligible improvements in classification accuracy across all scenes after introducing the incidence angle correction. The global strategy employed by IRGS effectively mitigates the effects of incidence angle variation, while the integration mechanism minimizes pixel-level errors

Table 2.9: Classification results before and after incidence angle (IA) correction for RF and IRGS-RF.

Scene ID	RF			IRGS-RF		
	Before IA correction	After IA correction	Difference	Before IA correction	After IA correction	Difference
20100418	87.60%	86.20%	-1.40%	86.10%	85.70%	-0.40%
20100426	88.00%	86.40%	-1.60%	88.10%	88.21%	0.00%
20100510	87.30%	88.30%	1.00%	88.00%	88.00%	0.60%
20100524	79.00%	77.40%	-1.60%	86.30%	86.30%	0.00%
20100623	90.80%	91.30%	0.50%	90.20%	90.18%	0.20%
20100629	71.70%	74.10%	2.40%	84.10%	84.27%	0.60%
20100721	87.40%	88.40%	1.00%	85.50%	85.34%	0.40%
20100730	73.70%	75.70%	2.00%	78.10%	78.23%	0.20%
20100807	93.30%	95.50%	2.20%	95.30%	95.28%	0.00%
20100816	86.80%	88.60%	1.80%	85.00%	84.91%	-0.60%
20100907	100.00%	99.80%	-0.20%	100.00%	100.00%	0.00%
20100909	98.80%	99.40%	0.60%	98.00%	97.68%	0.00%
20101003	77.20%	80.00%	2.80%	77.70%	77.70%	0.00%
20101021	90.80%	91.80%	1.00%	95.60%	95.72%	0.20%
20101027	72.20%	72.60%	0.40%	76.90%	76.90%	0.00%
20101114	68.80%	67.60%	-1.20%	75.60%	75.95%	-1.00%
20101206	83.00%	82.20%	-0.80%	86.60%	87.24%	-0.60%
20101214	72.70%	75.70%	3.00%	76.90%	76.90%	2.00%
Total	83.84%	84.50%	0.66%	86.33%	86.36%	0.09%

that might arise from outliers in the RF results. The classification results underscore the effectiveness of the proposed infrastructure in handling the complexities associated with the SAR image classification of sea ice.

The classification results indicate that GLCM features are qualified to distinguish different sea ice types and water when incorporated with a suitable classifier. Since SVM and RF employ different strategies to solve linearly separable problems, the separability of different sea ice types and the effectiveness of features that contributed to the classification results are worth investigating. Figure 2.5 shows the distributions of GLCM features for different sea ice types and open water. Open water can be efficiently separated from sea ice by some GLCM features, such as ASM and HOM. The distributions of COR for sea ice are highly correlated and are not feasible to discriminate sea ice types when utilized

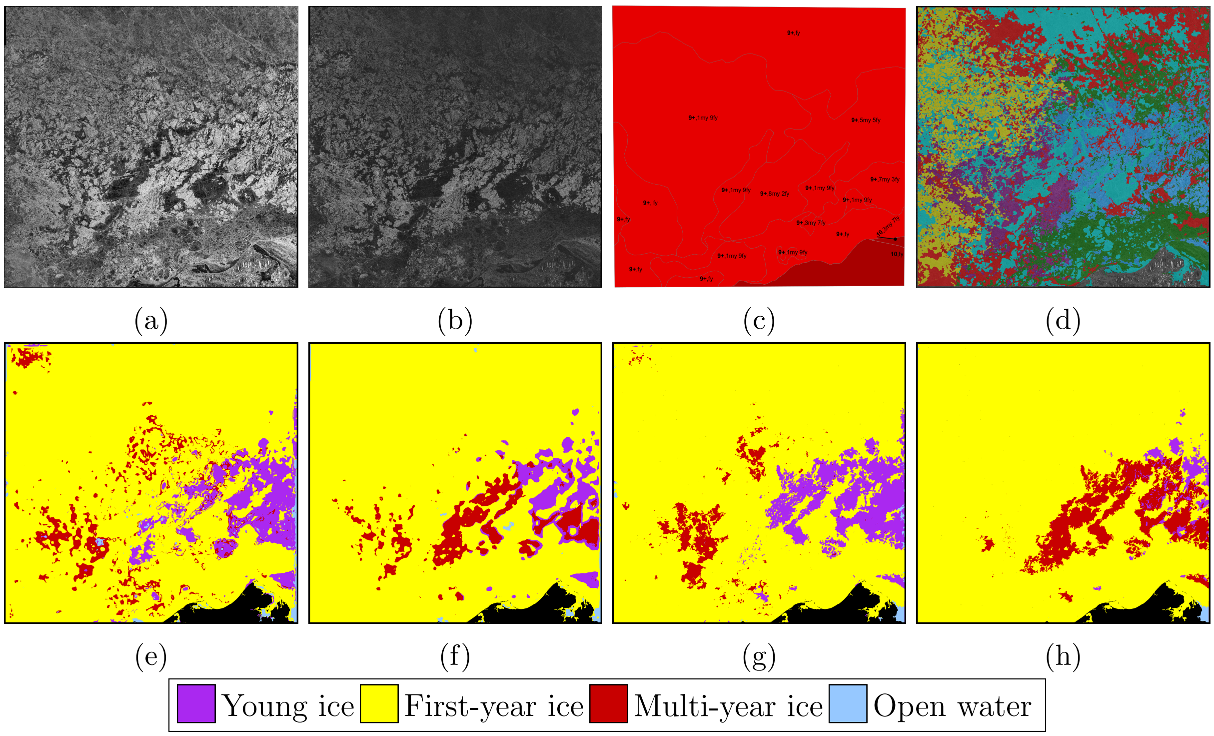


Figure 2.6: Classification results of April 26, 2010 (scene ID 20100426). (a) HH polarization. (b) HV polarization. (c) Ice chart. (d) IRGS segmentation result. (e) SVM pixel-based classification result with an accuracy of 83.40%. (f) RF pixel-based classification result with an accuracy of 88.00%. (g) IRGS-SVM classification result with an accuracy of 86.40%. (h) IRGS-RF classification result with an accuracy of 88.10%. According to the ice chart, there should not be YI in this scene. False YI is reduced by combining IRGS results. The final result is more reasonable.

solely. The distributions of HOM and INV are very similar, indicating information redundancy when using HOM and INV with the same parameters. In agreement with previous research, the results demonstrate that most GLCM features can contribute to the partial separation of open water and different sea ice types.

The classification results for the April 26, 2010 scene, which was previously shown in Figure 2.2 to demonstrate all steps of the proposed method, are shown in Figure 2.6. SVM has been demonstrated as an effective method to classify ice and water, but it does not generate a reasonable ice map in this scene. There are several issues that should be considered when applying pixel-based classifiers. First, the selection of crucial hyperparameters is essential. SVM requires tuned C and gamma, while the number of trees is the only key hyperparameter for RF. More hyperparameters require more computation for grid search and may cause overfitting. Second, SVM uses a kernel function to improve the separability of the features, and the speckle noise in the dataset is more likely to be augmented in the higher dimension. In contrast, the voting and bagging scheme in RF makes it less sensitive to speckle. Figure 2.6(e) displays noticeable noise-like errors across the whole scene. Unlike SVM, RF achieves more consistent classification results. According to visual interpretation and ice chart, there is no young ice in this scene. However, both SVM and RF misclassified first-year and multi-year ice in the middle right of the scene as young ice. On a pixel-wise classification level, RF obtains better accuracy compared to SVM. After being combined with the IRGS segmentation result, the noise-like errors are suppressed, and the boundaries are well preserved for both IRGS-SVM and IRGS-RF. Despite that, the improvement of IRGS-RF is negligible in numerical accuracy, because this is validated only on 500 sample points in each scene. The classified image using IRGS-RF is the most visually satisfying among the four methods. IRGS-RF also improves the misclassified young ice in the other three scenes. IRGS-RF also improves the misclassified young ice area.

An in-depth examination of a small region comprised of diverse sea ice types aids in gaining an improved understanding of the performance of various classification methods under complex sea ice conditions. Figure 2.7 presents a region of interest in Scene 20100426. In this scene, only FYI and MYI are present, as revealed by HH, HV polarized images, and ice chart data. These sea ice types are intertwined, rendering the region particularly challenging for sea ice classification due to its inherent complexity. FYI regions appearing at the bottom of the image exhibit a markedly different texture pattern from those appearing among MYI, amplifying the challenge of distinguishing them from MYI. Additionally, the disorderly and unsystematic intermingling of FYI and MYI in the image further complicates the classification task to determine boundaries between these two types of ice.

The pixel-wise methods, SVM and RF, encounter difficulties in accurately classifying MYI. However, RF shows a moderately better performance. SVM misidentifies the majority of MYI as YI, and the resulting ring-like artifacts around each ice region distort the natural boundaries. Moreover, the area on the right side of the image is misclassified as OW, indicating SVM's limitations in distinguishing water from ice using the extracted GLCM features.

The incorporation of segmentation results, through IRGS-SVM, manages to restore most ice boundaries and improve the accuracy of the misclassified MYI in the image. However, the overall classification performance remains suboptimal due to the initial poor results produced by SVM. Contrastingly, even though RF also delivers subpar pixel-level results, IRGS-RF demonstrates proficiency in correcting most of the misclassified areas and

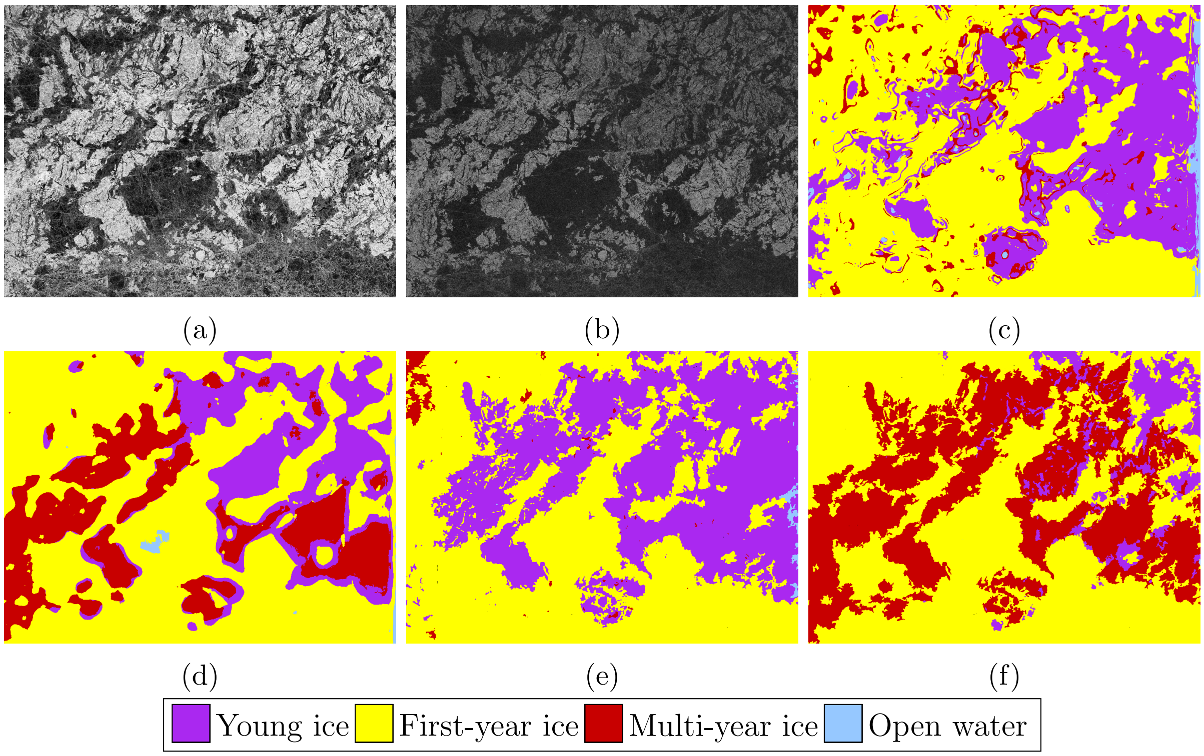


Figure 2.7: Classification results of a region of interest of April 26, 2010 (scene ID 20100426). (a) HH polarization. (b) HV polarization. (c) SVM pixel-based classification result. (d) RF pixel-based classification result. (e) IRGS-SVM classification result. (f) IRGS-RF classification result with.

re-establishing the natural boundaries among different sea ice types. These findings suggest that the integration with segmentation can refine classification accuracy and restore sea ice boundaries when pixel-level labels provided by SVM or RF are accurate.

However, the proposed method does not substantially improve the accuracy of initial pixel-level results if they are poor. In such instances, only sea ice boundaries can be restored. This highlights the criticality of obtaining accurate initial pixel-level labels for the successful application of our methodology.

Mapping sea ice during the summer melting season is usually tricky. The melting reduces the surface roughness of ice and degrades the texture captured by backscatter. The melting ponds presented onsite also change the electromagnetic characteristics of the ice beneath. An example scene obtained on August 07, 2010, is depicted in Figure 2.8. It

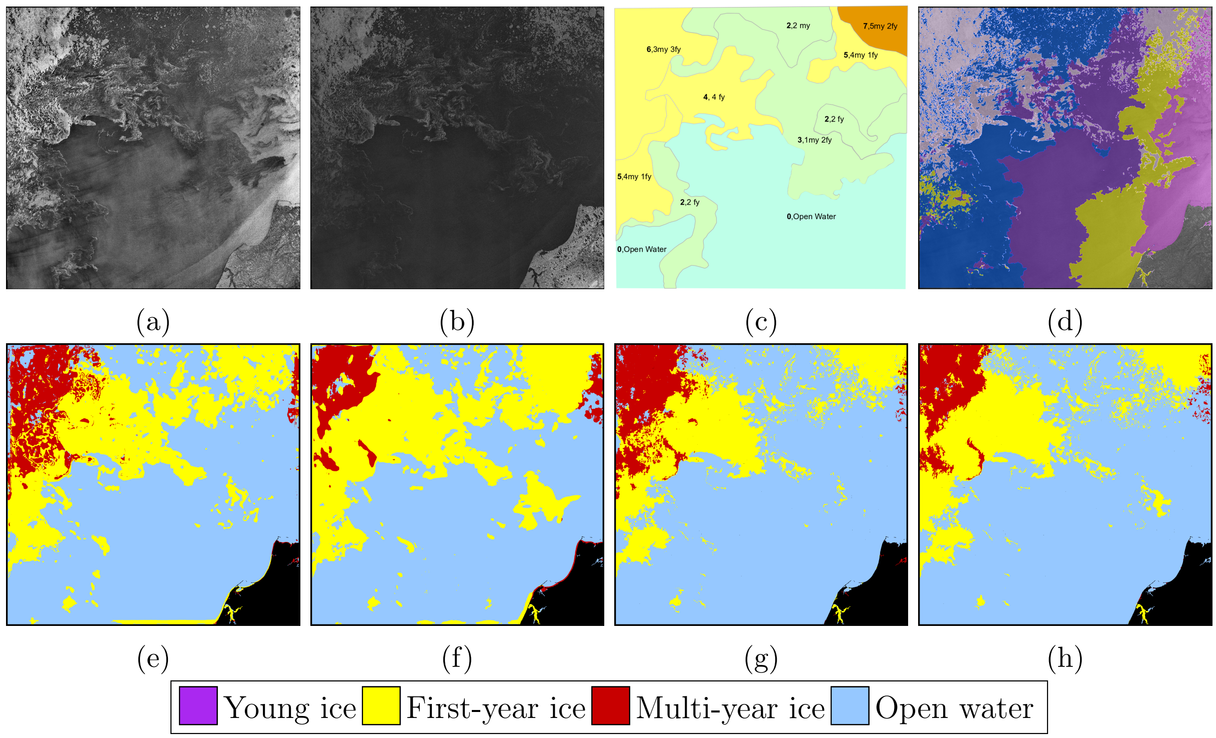


Figure 2.8: Classification results of August 7, 2010 (scene ID 20100807). (a) HH polarization. (b) HV polarization. (c) Ice chart. (d) IRGS segmentation result. (e) SVM pixel-based classification result with an accuracy of 91.90%. (f) RF pixel-based classification result with an accuracy of 93.30%. (g) IRGS-SVM classification result with an accuracy of 94.70%. (h) IRGS-RF classification result with an accuracy of 95.30%.

is a complex scene that contains first-year ice, multi-year ice, and open water. All four models distinguish the open water in the lower part of the scene. Although water around the image boundary is misclassified as first-year ice by pixel-wised SVM and RF, these errors are effectively mitigated by the combined models, IRGS-SVM and IRGS-RF. The upper part is more challenging since the surface roughness of sea ice is reduced by melting. Different types of ice show a very similar texture to open water in this scene, leading to confusion when discriminating between open water and sea ice types. Therefore, the presence of first-year ice is overstated by SVM and RF. The proposed IRGS-RF is robust to these first-year ice errors and achieves a higher classification accuracy of 95.30%. Most multi-year ice floes in the upper right corner are also preserved in the results.

The scene of September 07, 2010 (Figure 2.9) is the least challenging scene in the

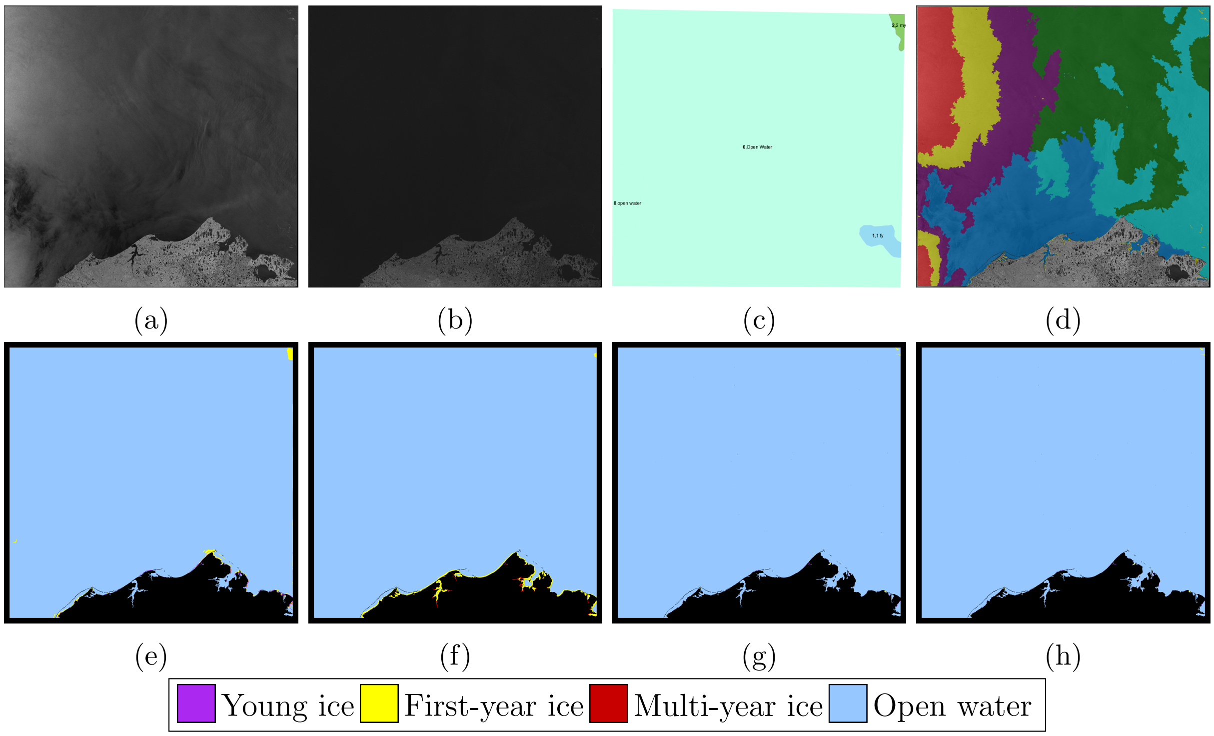


Figure 2.9: Classification results of September 07, 2010 (scene ID 20100907). (a) HH polarization. (b) HV polarization. (c) Ice chart. (d) IRGS segmentation result. (e) SVM pixel-based classification result with an accuracy of 98.50%. (f) RF pixel-based classification result with an accuracy of 100.00%. (g) IRGS-SVM classification result with an accuracy of 100.00%. (h) IRGS-RF classification result with an accuracy of 100.00%.

dataset. The scene is acquired at the end of summer and is only covered by water. The texture appears in the right of the scene, and dark areas in the bottom left corner are caused by strong wind roughening. Both SVM and RF misclassified some in-land and near-land areas. Since pixels near land and boundary are discarded for numerical validation, RF still achieves 100% accuracy on the test pixels. After combining with segmentation, all these misclassified pixels are eliminated. The nearly perfect sea ice maps obtained by IRGS-SVM and IRGS-RF are presented in Figure 2.9(g) and 2.9(h).

The classification result of October 21, 2010 is displayed in Figure 2.10. The grey ice appearing in the top left corner has a much lower backscattering level than other grey ice displayed in the scene. Another challenge presented is the noticeable banding artifacts in the middle of the HV scene. The dataset is collected under ScanSAR wide beam mode,

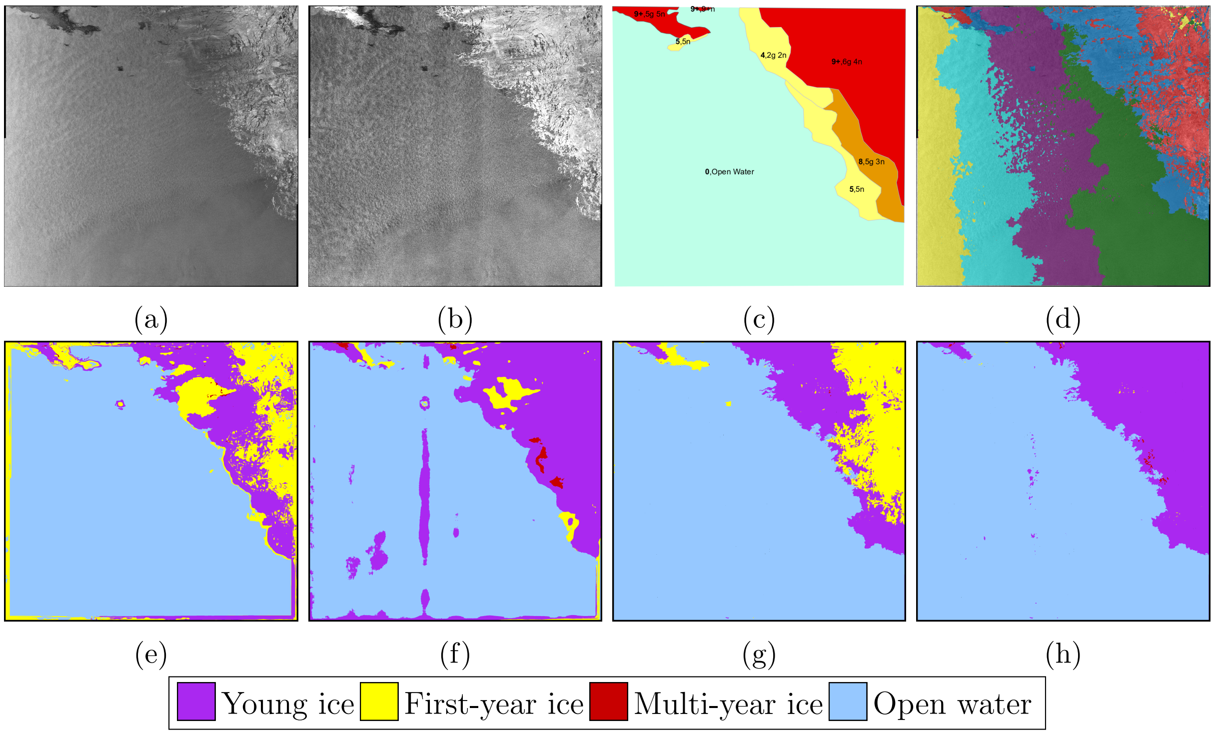


Figure 2.10: Classification results of October 21, 2010 (scene ID 20101021). (a) HH polarization. (b) HV polarization. (c) Ice chart. (d) IRGS segmentation result. (e) SVM pixel-based classification result with an accuracy of 89.40%. (f) RF pixel-based classification result with an accuracy of 90.80%. (g) IRGS-SVM classification result with an accuracy of 92.60%. (h) IRGS-RF classification result with an accuracy of 95.60%.

and adjoining multiple scan beams cause these vertical bands. Since HV polarization has a much lower signal-to-noise ratio, the banding artifacts are quite common in the HV scene. SVM and IRGS-SVM do not get affected by banding artifacts. However, there is no first-year ice in this scene, and SVM and IRGS-SVM misclassify young ice as first-year ice. RF obtains much better results than SVM. Although the banding artifacts greatly impact the pixel-wise RF result, the combination with segmentation almost resolves this issue, leaving only a small error on the IRGS-RF result.

Figure 2.11 (November 14, 2010) represents the most challenging scene with different ice types and water. The proposed IRGS+RF achieves the lowest accuracy of 75.60% of all 18 scenes. First, the incidence angle effect is significant in the scene. The backscattering signature is inconsistent across the whole image–the left of the scene is brighter than the

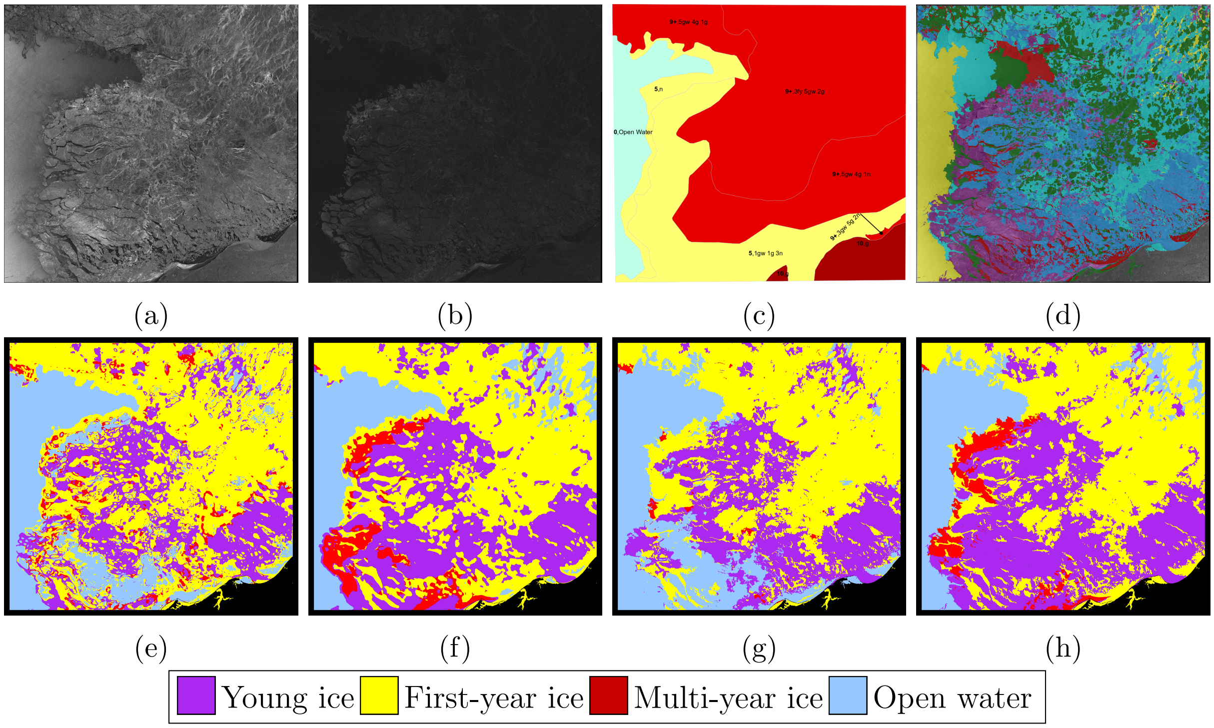


Figure 2.11: Classification results of November 14, 2010 (scene ID 20101114). (a) HH polarization. (b) Ice chart. (c) Ice chart. (d) IRGS segmentation result. (e) SVM pixel-based classification result with an accuracy of 60.50%. (f) RF pixel-based classification result with an accuracy of 68.80%. (g) IRGS-SVM classification result with an accuracy of 65.40%. (h) IRGS-RF classification result with an accuracy of 75.60%.

right side. Second, given the 200 m X 200 m pixel size, a single pixel may contain several types of ice and water. The pixel-wise classifiers struggle to distinguish these mixed pixels. Third, there are numerous leads presented among sea ice in this scene. These leads may have generated misleading texture features that confused the classifiers. Although the pixel-wise results are unsatisfying, the combined methods suppress these phenomena and achieve higher classification accuracy and natural boundaries between different ice types and water.

The experiments are run on a computer with the following configuration: Intel Core i5-6600K, 16-GB RAM, and Windows 10 operating system. The average execution time to generate a sea ice map based on a RADARSAT-2 scene is less than 25 min. Specifically, it takes 3 min to oversegment the scene into homogeneous regions. The GLCM feature extraction, as the most time-consuming part of the workflow, takes around 15 min. SVM takes 5 min for pixel-wise labeling, while random forest only takes 20 seconds credited to parallel computing. The proposed system is qualified to deploy on business computers with average configurations and classify sea ice in a scene within half an hour.

2.6 Summary

An automatic sea ice classification infrastructure using RADARSAT-2 SAR imagery is proposed in this article. To the authors' best knowledge, this is the first study combining segmentation with pixel-wise labeling, using an energy function, to classify different sea-ice types. The unsupervised IRGS segmentation algorithm extracts spatial contextual information in the SAR scene to divide the whole image into homogeneous regions, while the pixel-wise classifier exploits backscatter intensities and textural features to label each region. Two benchmark pixel-wise classifiers, SVM and RF, and two proposed models, IRGS-SVM and IRGS-RF, were trained and tested on a dataset to find the best combination for building the system.

To better evaluate the proposed models, a dataset consisting of 18 RADARSAT-2 scenes of the Beaufort Sea is used to evaluate the proposed models. The dataset includes scenes from melting, summer, and freezing seasons. The LOO strategy is applied for cross-validation to avoid using samples from the same scene for both training and testing. The results show that the proposed models achieve an overall accuracy of 86.33% on the dataset and are robust to melting season, which is the most challenging period of the year.

When only applying pixel-wise classifiers, RF obtains an overall accuracy of 84.07% compared to 81.13% by SVM. Comparing the visual results, the sea ice maps generated by RF contain fewer noise-like errors than SVM. In general, RF outperforms SVM on most of the scenes in the dataset, indicating that RF is a more reliable choice when dealing with sea-ice classification based on texture features. After combining IRGS segmentation results with the pixel-level labels, the classification accuracies of IRGS-SVM and IRGS-RF are both improved. IRGS-RF achieves the best performance with an 86.33% success rate for distinguishing ice types and water.

Additionally, we have conducted experiments on a normalized dataset to assess the robustness of the proposed method to variations in incidence angle. The classification accuracies of the RF classifier showed improvement for the majority of scenes when the SAR imagery was normalized to an incidence angle of 35°. Nevertheless, the IRGS-RF demonstrated remarkable stability, yielding almost identical overall accuracies on both the normalized and unnormalized datasets.

This consistency in performance can be attributed to the effectiveness of the unsupervised segmentation in capturing contextual information from each homogeneous region, subsequently mitigating the classification errors in the RF results. It is worth noting that such robustness is crucial, especially when dealing with dual-pol SAR data, as it ensures the reliability and accuracy of the classification results despite variations in incidence angle. The results indicate that the proposed method, with its inherent ability to handle incidence angle variation, stands as a robust solution for sea ice classification in dual-pol SAR imagery.

The novel sea ice classification approach presented in this work not only delivers promising results in terms of accuracy but also generates visually appealing ice maps. It successfully preserves the natural boundaries between water and various types of ice while refining pixel-level inaccuracies. The resultant sea ice maps exhibit a high level of consistency with the ice charts provided by CIS, making them valuable references for sea ice interpretation within operational contexts at ice services. These outcomes support the assertion that the proposed methods effectively merge both textural and contextual features, significantly enhancing the precision of sea ice classification.

In future endeavors, the proposed model will undergo rigorous testing on datasets that exhibit diverse temporal and textural characteristics, utilizing transfer learning to optimize its performance. Additionally, the impact of applying noise floor correction will be analyzed to refine the model's accuracy further. Most crucially, while the current approach integrates contextual and textural features within individual regions derived from segmentation results, it does not account for the global context in SAR images. Future research will thus be directed towards developing a new methodology capable of incorporating this global context into the sea ice classification process, with the aim of achieving even more accurate and reliable results.

Chapter 3

Sea ice–water Classification of RADARSAT-2 Imagery Based on Residual Convolutional Neural Network with Regional Pooling

Sea ice mapping plays an integral role in ship navigation and meteorological modeling in the polar regions. Numerous published studies in sea ice classification using synthetic aperture radar (SAR) have reported high classification rates. However, many of these focus on numerical results based on sample points and ignore the quality of the inferred sea ice maps. We have designed and implemented a novel SAR sea ice classification algorithm where the spatial context, obtained by the unsupervised IRGS segmentation algorithm, is integrated with texture features extracted by a residual convolutional neural network and, using regional pooling, classifies ice and water. This algorithm is trained and tested on a published dataset and cross-validated using a leave-one-out (LOO) strategy, obtaining an overall accuracy of 99.67% and outperforming several existing algorithms. In addition, visual results show that this new method produces sea ice maps with natural ice–water boundaries and fewer ice and water errors.

3.1 Introduction

Sea ice covers about 12% of the oceans on Earth [89]. In high latitude and polar regions, sea ice reduces the heat exchange between the sea and the atmosphere, regulating the global climate [8]. As the global temperature has been rising in the past decades, sea ice thickness has reduced dramatically [90]. Melting ice poses a significant impact on the ecosystem and meteorology in the Arctic region. Meanwhile, exploring shipping routes and marine resources becomes attractive in the summer season [91]; therefore, monitoring sea ice distribution and how it changes in the life span is essential.

Synthetic aperture radar (SAR) [92] is a reliable method to monitor sea ice because SAR imagery can be acquired day and night under any type of weather condition. Popular satellites deployed for analyzing sea ice are the Sentinel-1 mission (operated by the European Space Agency) and the RADARSAT system (RADARSAT-2 and RADARSAT Constellation Mission (RCM), operated by the Canadian Space Agency). Ice agencies from different nations, e.g., the Canadian Ice Service (CIS), the National Ice Agency, and the Norwegian Ice Service, process SAR data and produce ice charts manually. With the expanding data volume received from recently launched SAR satellites [93, 94], the demand for automated sea ice classification systems is growing.

A typical sea ice classification system usually consists of two parts. First, handcrafted features are extracted from each pixel. Second, an appropriate classifier is trained on the feature set and predicts each pixel's label in the scene. For the first step, originally only backscattering intensities were used to distinguish sea ice and water [16]; however, the non-stationarity caused by weather conditions (e.g., wind speed and melting ponds on ice surface [95, 96, 97]) and satellite's infrastructure (e.g., incidence angle effect [98, 99, 31] and speckle noise [100, 101, 102]), make it not possible to solely use backscatter for operational sea ice classification tasks. Meanwhile, studies [18, 103, 104] reported high classification accuracy by utilizing polarimetric features. Gill and Yackel [17] extracted polarimetric parameters from quad-polarized RADARSAT-2 imagery using different decomposition methods. K-means and maximum likelihood classifiers were adopted to discriminate different types of first-year ice. Although quad-polarized SAR imagery shows enormous potential in classifying sea ice, the narrow swath, which is on the order of 50 km, does not have sufficient coverage for operational sea ice monitoring that utilizes swaths in the range 400–500 km. Dabboor et al. [47] trained a random forest (RF) classifier using 23 simulated compact polarimetric (CP) features extracted from quad-polarimetric (QP) data to classify first-year and multi-year ice. Ghanbari et al. [105] also used simulated CP features derived from QP data to classify different ice types. Even though the classification accuracy achieved by CP features is lower than that of QP features, the wider swath (350 km for RCM) of CP data makes it a better data source for operational sea ice monitoring.

Since major national sea ice agencies favor imagery with large area coverage for operational use, dual-polarization imagery with a 500 km swath has become the main data source for sea ice mapping in the past decade [72]. In addition to backscatter intensities, textural features extracted using the gray level co-occurrence matrix (GLCM) [106] can be used to enhance the feature description for sea ice discrimination [23, 107]. Clausi [24] explored the classification performance using different GLCM measurements. The study showed that the classification accuracy was not always rising with the increasing grey levels of quantization. Li et al. [108] proposed an unsupervised method to classify ice and water. The HV scene was segmented into homogeneous regions using a modified watershed algorithm. Then, the Otsu threshold was applied to distinguish the homogeneous regions, which were chosen as training samples. A support vector machine was trained on GLCM features and tested on 728 Sentinel-1 extra-wide images. Lyu et al. [109] extracted GLCM features and trained a random forest to separate sea ice from water in RCM data.

Deep learning has recently been introduced to remote sensing because of its phenomenal achievement in the computer vision domain. The classification results that use features learned by deep learning models are usually comparable, sometimes superior to those using traditional engineered features [110]. The convolutional neural network (CNN), a particular deep learning structure, has been widely adopted in sea ice classification recently [111, 112, 41, 113, 40]. Ren et al. [114] proposed a two-step deep learning model named (DAU-Net) to discriminate between sea ice and open water. A residual neural network (ResNet) was deployed to extract features from input SAR imagery. Then, a fully connected U-Net integrated with a dual-attention mechanism ingested the learned featured map and produced an ice–water classification result. The model was trained on 15 dual-polarized scenes and tested on the other three. The DAU-Net improved the intersection over union (IoU) compared with the original U-Net. Junhwa et al. [115] used long- and short-term memory (LSTM) to capture temporal relation between SAR images. A deep learning model consisting of encoders, LSTMs, and decoders was developed to predict sea ice. The model used a novel perceptual loss function and accurately predicted sea ice concentration.

However, most deep-learning-based studies compare their methods with benchmark deep-learning models or traditional classifiers with the input of backscattering intensity. The advantages of using the learned feature (deep learning) compared with engineered features (e.g., GLCM) have not been sufficiently investigated. Moreover, boundaries between sea ice and water, which are well presented in ice charts, are usually corroded in the classification results [105]. Since pixel-level ground truth is scarce, the performance of the sea ice classification method is generally trained and evaluated based on sample points rather than the whole scene. Many studies prefer to select samples from regions with high concentrations and no boundaries to ensure the quality and quantity of the training data. The classifiers can not learn the characteristics of the ice–water boundary based on limited samples. Furthermore, both CNN- and GLCM-based methods extract features using sliding windows, and in each window, all pixels inside contribute to describing the features of the center pixel. Figure 3.1 depicts an extreme case for learning window-based features. It is a 9-by-9 image path used to extract the feature of the center pixel “X”. “A”, “B”, “C”, and “D” represent the other four classes. Although the path is designed to learn features of pixel “X”, it extracts the features derived from its neighbor, class “A”, “B”, “C”, and “D”, rather than itself. The features derived from mixed classes confuse the classifier and lead to classification errors [44].

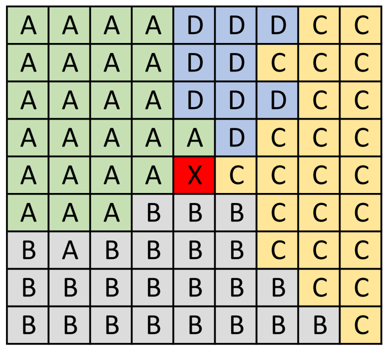


Figure 3.1: Example of an image patch for feature extraction. “X” is the center pixel that needs to be extracted features. “A”, “B”, “C”, and “D” are four different classes.

Considering the challenges associated with applying deep learning models directly to SAR imagery for sea ice classification, this study aims to explore efficient methods for learning spatial features specific to sea ice in SAR images, and how to integrate these spatial features with contextual information gleaned from semantic segmentation algorithms. However, due to the scarcity of labeled data for different types of sea ice, this research will primarily concentrate on differentiating sea ice from open water. Residual neural networks, which have demonstrated considerable success in remote sensing applications, are employed to capture the spatial characteristics of sea ice and water. Concurrently, Iterative Region Growing with Semantics (IRGS) [70]–a semantic segmentation algorithm particularly crafted for remote sensing data—is deployed to delineate the ice–water boundary by extracting contextual features. These two components are synthesized through an innovative region-pooling layer designed to enhance the classification of sea ice in SAR imagery. The following are the main contributions of this work.

1. We propose a novel end-to-end sea ice–water classification system based on a deep learning model using SAR imagery. One of the major attractions of the proposed system is that it can generate a pixel-level classification result while the fine boundaries between ice and water are well preserved.

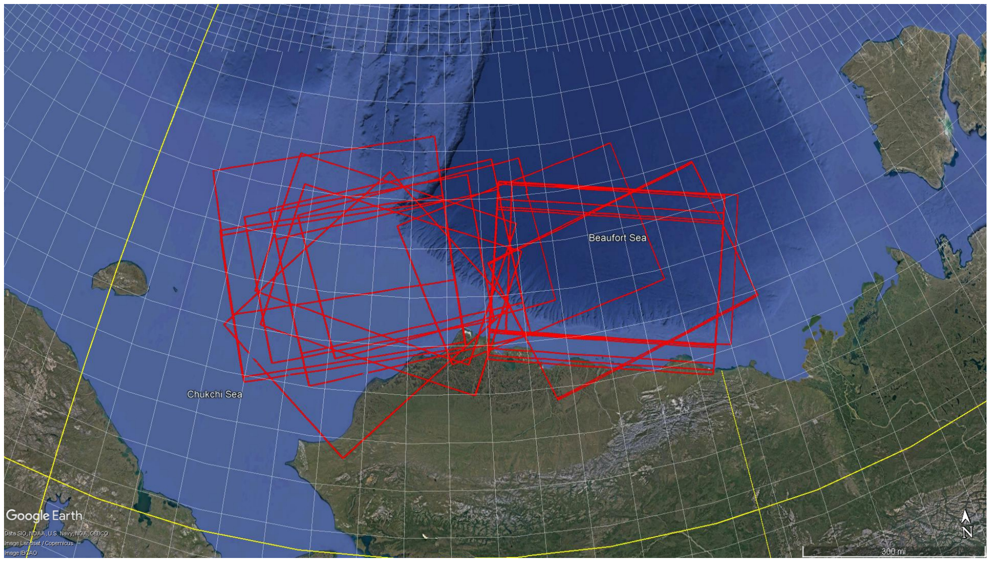


Figure 3.2: Location of the Beaufort Sea. Footprints of the 21 RADARSAT-2 scenes used in this work are shown in yellow.

2. We conduct a comprehensive set of experiments to compare the effectiveness of engineered features versus model-learned features in the classification of sea ice and open water. The findings reveal that model-learned features surpass their engineered counterparts, obviating the need for human expertise in feature engineering.

3. We explore the classification capability of a deep learning model with different input and patch sizes. The results obtained by the deep learning model with different hyper-parameters provide a baseline reference for future work.

4. We extensively evaluate the performance of the proposed model and compare it with two benchmark methods and two reference methods. The results show that our model outperforms these methods of comparison both numerically and visually.

3.2 Data

3.2.1 The RADARSAT-2 ScanSAR Wide Mode Dataset

The dataset used in this paper to train and validate the proposed method contains 21 scenes at different locations over the Beaufort Sea in 2010. The Beaufort Sea can be

Table 3.1: List of SAR scenes used in this work.

Scene ID	SAR Acquisition Date (M/D/Y)	Acquisition Time UTC (hh:mm:ss)	Ascending (A)/ Descending (D)	Incidence Angle Near Range (°)	Incidence Angle Far Range (°)
20100418_163315	18 April 2010	16:33:16	Descending	19.72	49.46
20100426_040439	16 April 2010	04:04:39	Ascending	19.55	49.44
20100510_035620	10 May 2010	03:56:20	Ascending	19.58	49.44
20100524_034756	24 May 2010	03:47:56	Ascending	19.61	49.46
20100605_163323	5 June 2010	16:33:23	Descending	19.77	49.46
20100623_041255	23 June 2010	04:12:55	Ascending	19.63	49.45
20100629_163326	29 June 2010	16:33:26	Descending	19.71	49.43
20100712_031834	12 July 2010	03:18:34	Ascending	19.61	49.39
20100721_173208	21 July 2010	17:32:08	Descending	19.64	49.47
20100730_162908	30 July 2010	16:29:08	Descending	19.61	49.46
20100807_173610	7 August 2010	17:36:10	Descending	19.77	49.47
20100816_163329	16 August 2010	16:33:29	Descending	19.74	49.45
20100907_035614	7 September 2010	03:56:14	Ascending	19.59	49.44
20100909_163321	9 September 2010	16:33:21	Descending	19.63	49.48
20101003_163324	3 October 2010	16:33:24	Descending	19.59	49.46
20101021_041325	21 October 2010	04:13:25	Ascending	19.50	49.43
20101027_025726	27 October 2010	02:57:26	Ascending	19.58	49.43
20101114_041304	14 November 2010	04:13:04	Ascending	19.57	49.43
20101120_163324	20 November 2010	16:33:24	Descending	19.70	49.44
20101206_015139	6 December 2010	01:51:39	Ascending	19.59	49.39
20101214_025725	14 December 2010	02:57:25	Ascending	19.58	49.45

considered as a marginal sea of the Arctic Ocean located north of Canada and Alaska. Figure 3.2 shows the geographical distribution of the dataset. Each scene was captured under ScanSAR wide beam mode and consisted of HH and HV polarizations from both ascending and descending satellite passes. Table 3.1 lists the scene ID, acquisition time, ascending or descending orbit, and incidence angle range of the scenes in the dataset. The average image size is around 10,500 by 10,000 pixels with a spatial resolution of 50 by 50 m. The nominal swath width is 500 km in both range and azimuth direction, and the incidence angle varies from 19.50 to 49.48 degrees. The 21 scenes were acquired from April through December under various ice–water conditions. Canadian Ice Service (CIS) acquired these images for manual interpretation and generating ice charts in 2010. This is the same dataset used by Leigh et al. [50] and Jiang et al. [116].

An example scene of the dataset, which was taken on 24 May 2010, is shown in Figure 3.3. This scene contains first-year ice, multi-year ice, water, and land. Both HH and HV are displayed, and HV is less sensitive to the incidence angle effect compared with HH. The first-year ice presented at the top of the scene appears visually different than that at the bottom. The water shows decreasing backscattering in the horizontal direction in the scene. There is another open water area that appears at the bottom right corner.

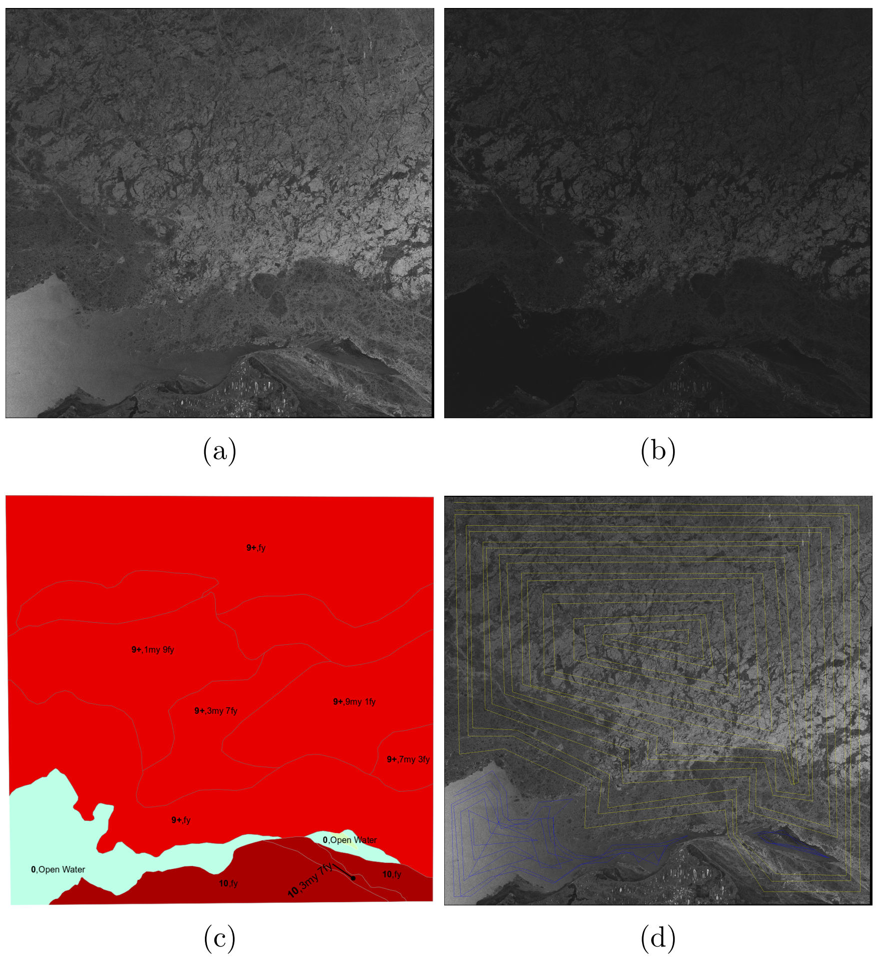


Figure 3.3: An example scene in the dataset. The scene was acquired on 24 May 2020. (**a**) HH polarization. (**b**) HV polarization. (**c**) Ice chart. (**d**) Sample points on poly-lines used for training and testing. ice: yellow, water: blue.

3.2.2 Data Pre-Processing

The pre-processing of the dataset used in this study includes radiometric calibration, down-sampling, and normalization. The first step is a fundamental processing to convert raw digital numbers (DN) to actual backscatter received by the sensor, ensuring the backscatter in SAR images captured from different time and locations are consistent and comparable, which is crucial for classification tasks in remote sensing. Down-sampling helps to reduce the data volume and improve the runtime of classification algorithms, while normalization is required by machine learning methods to facilitate

The product package supplied by MDA Ltd. comes inclusive of lookup tables (LUTs) to facilitate straightforward transformation from raw DN to backscatter coefficients including sigma-nought (σ^0), beta-nought (β^0), and gamma (γ). For this study, we engaged the offset and gain values specified in the sigma-nought LUT to convert the raw data into σ^0. The conversion process adheres to the formula as defined in[117], which is elucidated below.

$$\sigma^0 = 10log\left(a_2\left(d^2 - a_1 n(r)\right) + a_3\right) \tag{3.1}$$

where d is the pixel intensity with a range from 0 to 255. a_1, a_2, and a_3 are noise scaling, linear conversion, and offset, respectively. $n(r)$ is the noise as a function of range r. Sigma nought is the calibrated backscatter coefficient and expressed decibels (dBs). Unlike passive sensors, images acquired by SAR sensors are usually contaminated by a multiplicative noise called speckle noise. Speckle noise is caused by the infrastructure of the SAR platform and interferes with the backscatter captured in SAR imagery [118]. Many studies use filters, e.g., the Lee filter, to remove speckle noise; however, applying the Lee filter does degrade the ice–water boundary information, which is crucial for sea ice–water classification. Hence, no speckle noise filter is utilized in this study.

Each image is approximately 10,500 by 10,000 pixels with a nominal pixel spacing of 50×50 m. A 4-by-4 average pooling is applied to the images in the dataset to reduce computational cost and runtime. The downsampled image size is around 2600 by 2500. The new 200 m pixel size is still adequate for producing sea ice classification maps with far more details than the ice charts interpreted by human experts [50].

Compared with HH, HV imagery usually has a much lower signal-to-noise ratio. Since the output of each layer in a deep learning model depends on the input value, the gradient descent trends to update some weights much faster than others if the scales of input channels vary; therefore, normalizing input features to a similar scale helps accelerate learning speed and produce faster convergence. For training and testing the proposed deep learning model, the backscattering intensities of all input scenes are normalized to the range of 0 to 1 using the corresponding minimum and maximum [40].

3.2.3 Dataset for Training and Validation

Pixel-level ground truth is important for training supervised models; however, one of the biggest challenges in sea ice classification is the lack of reliable labeled samples. The most common reference data source is the ice chart released by CIS. Figure 3.3(c) shows an ice chart over the Beaufort Sea from 24 May 2010. There are two steps involved in producing sea ice charts. First, the trained operators break down the scene into defined regions called 'polygons'. Then, each polygon is interpreted based on the ice concentration and visually recognized sea ice types. For example, the text "9+, 1my, 9fy" on the polygon indicates that the overall ice concentration in this polygon is more than 90%, with 10% multi-year ice and 90% first-year ice for ice coverage.

To guarantee the integrity and reliability of the training data, we engaged the expertise of a former ice analyst from CIS to create detailed ice charts for our dataset. Given the nature of ResNet, which requires an extensive amount of labeled data to learn and distinguish between sea ice and open water accurately, we made our best efforts to generate labeled pixel samples using the aforementioned ice charts as reference.

The initial step involved densely drawing polylines within homogeneous regions. Pixels situated on the yellow polylines were designated as ice samples, whereas pixels on the blue polylines were classified as water samples. We ensured that polylines spanned across the entire image, deliberately avoiding regions close to ice–water boundaries to prevent the inclusion of pixels with ambiguous labels in the labeled dataset.

It is important to note that the sample pixels forming the polylines exhibit a high degree of correlation, a factor that could potentially lead to overfitting. To mitigate this risk, we decided to sample only 20% of the pixels along the polylines sparsely, instead of utilizing every pixel, for the construction of the labeled dataset. This approach not only proved to be efficient but also helped in reducing spatial correlation, thereby enhancing the generalizability of the classification model. A comprehensive summary of the labeled samples is presented in Table 3.2. In scenarios where a scene is dominated by a single class, the sample count for the other classes is recorded as zero.

Although using a dataset that contains hundreds of scenes to evaluate the method's performance is attractive, some studies [108, 52] only employ a small portion of the dataset for numerical testing due to the lack of detailed ground truth. Moreover, the results in the article published by Zhang et al. [48] indicate that the classification accuracy follows

Table 3.2: Number of labeled samples used in this study.

Scene ID	# of Water	# of Ice
20100418_163315	0	38,323
20100426_040439	0	49,484
20100510_035620	3,071	54,899
20100524_034756	9,307	50,780
20100605_163323	1,388	52,588
20100623_041255	21,569	42,362
20100629_163326	6,564	26,190
20100712_031834	5,072	24,702
20100721_173208	10,002	12,638
20100730_162908	9,919	8,551
20100807_173610	9,092	2,680
20100816_163329	12,866	10,689
20100907_035614	24,201	0
20100909_163321	23,094	1,598
20101003_163324	18,605	8,414
20101021_041325	25,138	8,554
20101027_025726	13,230	14,334
20101114_041304	8,219	13,597
20101120_163324	0	47,944
20101206_015139	0	47,850
20101214_025725	0	40,658
Total	201,338	556,837

a similar distribution for consecutive years. Since our dataset covers a whole year of sea ice life span, it is sufficient to assess the proposed method [119].

The prevalent evaluation criterion in most studies on sea ice classification is classification accuracy. A customary practice involves using the same scene for both training and testing, or employing a shuffle method to partition the training dataset into distinct sections for training, validation, and testing. While this approach tends to inflate the classification accuracy, due to the shared characteristics of data within the same scene, it also frequently results in reduced accuracy in operational tasks. This is largely attributable to the variability in the backscattering of sea ice and water across different scenes.

In order to ensure the robustness and generalizability of IceNet, we propose an evaluation scheme based on leave-one-out cross-validation [120]. The composition of our dataset, encompassing 21 separate scenes, offers a unique opportunity to structure the evaluation procedure in a way that maximizes the usefulness of our data. Initially, the process involves partitioning the original samples into 21 subsets. Each subset exclusively comprises samples from a single scene. For every iteration, one subset is retained as the test data for model evaluation. The remaining 20 subsets form the training and validation data, maintaining a training/validation split ratio of 0.7. IceNet is then trained and validated on these 20 images and evaluated on the left-out image. This process is repeated such that each image in the dataset is used once as the test data. This process helps to avoid overfitting, ensures that every sample contributes to the evaluation of the model, and provides a comprehensive assessment of IceNet's performance. Finally, the overall performance metric is calculated by averaging the results across all 21 iterations. This approach allows us to obtain a more realistic, reliable, and generalizable measure of IceNet's performance.

3.3 Method

The architecture of the proposed sea ice classification model consists of two main parts: the iterative region growing with semantics (IRGS) segmentation and residual CNN labeling. Figure 3.4 shows the framework of the model, which is named as IceNet. The input HH and HV images are first processed for radiometric calibration and down-sampling. Then, the contextual information is extracted by IRGS using the pre-processed images, while the spatial features are learned by residual CNN. Finally, the contextual and spatial information are combined based on a novel regional pooling layer.

3.3.1 Unsupervised Model for Segmentation

Markov Random Fields (MRFs) and Conditional Random Fields (CRFs) have gained popularity in image segmentation due to their capability to model contextual information

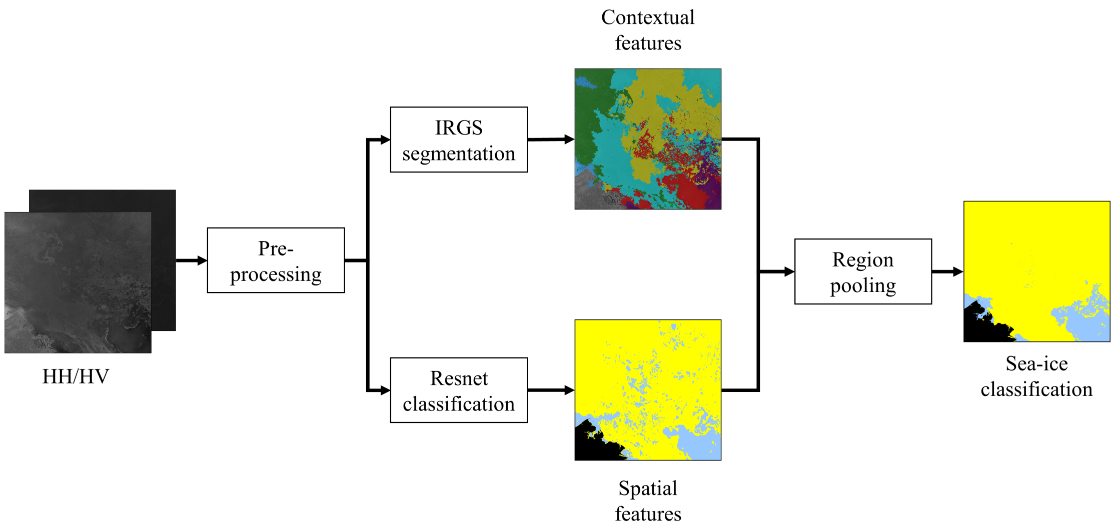


Figure 3.4: Flow diagram of the proposed IceNet for sea ice classification.

within images effectively. The IRGS algorithm, grounded in MRF principles, has been developed to offer a robust segmentation method tailored for remote sensing imagery. In this research, we have applied IRGS, integrating it with a 'glocal' strategy [50], aiming to learn contextual information that helps to preserve the boundaries between ice and water. Figure 3.5 elucidates the fundamental steps involved in the IRGS process.

The initial phase of the process involves segmenting the scene into smaller regions, referred to as autopolygons. Given that a lower signal-to-noise ratio characterizes HV polarization and demonstrates reduced sensitivity to variations in incidence angle compared to HH polarization, we opted to utilize solely the HV polarization in the generation of autopolygons. This decision was motivated by the intention to minimize distortion caused by incidence angle variation on the statistics of the backscatter coefficient.

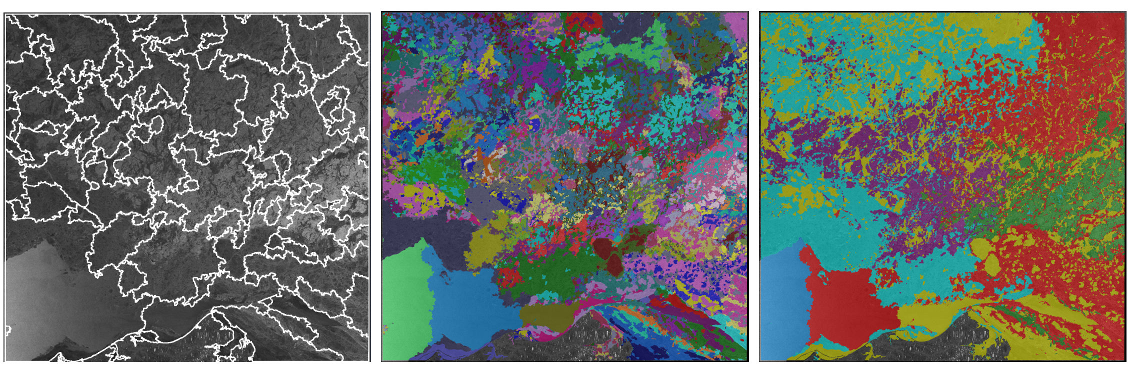
The process commences with the division of the scene into a 12-by-12 grid net, wherein the center pixel within each grid serves as a seed. Subsequently, a watershed algorithm is applied to the HV polarization, guided by these seeds. The autopolygons yielded from this watershed process are depicted in Figure 3.5(a). Following this, the IRGS segmentation is executed on each autopolygon, with both HH and HV scenes serving as input.

Within the IRGS pipeline, a modified watershed algorithm is employed to oversegment each autopolygon. This oversegmentation then serves as the basis for constructing a Region Adjacent Graph (RAG). The design of the RAG aims to minimize the following energy function:

$$E = \sum_{i \in R} V_G(x_i) + \sum_{\langle i,j \rangle \xi} V_E(x_i, x_j) \tag{3.2}$$

Here, $V_G()$ represents the unary potential, while $V_E()$ signifies the pairwise potential in the MRF models. Within the context of IRGS, $V_G()$ corresponds to Gaussian statistics for regions derived from oversegmentation, and $V_E()$ accounts for edge strength between cliques ξ (connected regions). Figure 3.5(b) showcases the result of this segmentation step. The limited size of the autopolygons ensures a constrained incidence angle variation, facilitating improved extraction of contextual information.

Post application of IRGS to all individual autopolygons, a new RAG is formulated over the entire HH and HV scenes. The objective here is to minimize the energy function on a global scale, culminating in producing the final segmentation results, as illustrated in Figure 3.5(c). This strategy, whereby the energy function is minimized locally (within each autopolygon) prior to a global minimization (across the entire scene), has been termed 'glocal'. Such an approach ensures that accurate local contextual information is first extracted, subsequently contributing to the formation of a coherent global context.



(a) (b) (c)

Figure 3.5: Steps of IRGS using the glocal strategy for the scene from 24 May 2010 (scene ID 20100524_034756). (**a**) Autopolygons overlays on the HH scene. (**b**) Local segmentation result. (**c**) Glocal segmentation result.

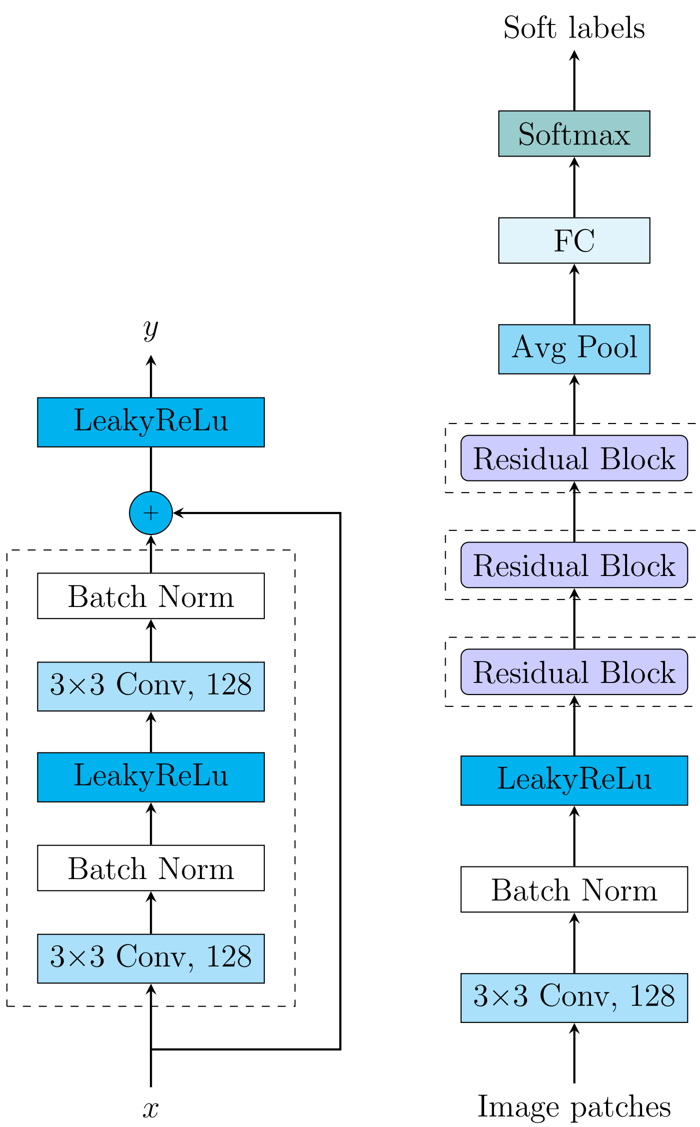
3.3.2 Deep Learning Model for Labeling

A typical CNN consists of the input layer, hidden layer(s), and output layer [121]. The output of each layer is the input of the next layer for both forward and backward propagation. Unlike conventional artificial neural networks, CNN has multiple types of hidden layers, including convolutional (ConV) layers, activation layers, pooling layers, and fully connected (FC) layers. The convolutional layers extract features from the input using multiple kernels of different sizes. Then, the learned features pass through activation and pooling layers for non-linearization and compression. The fully connected layer maps all learned high-level features to sample space before the final classification.

However, as the network's depth goes deeper, the propagation of gradients through numerous layers during backpropagation can lead to challenges in training stability and convergence. When the gradients of the loss are very large, it can result in excessively large updates to the weights, destabilizing the network. This issue is known as the exploding gradient problem. Conversely, when the gradients are very small, they can diminish as they are backpropagated through the network, approaching zero. This results in minimal updates to the weights, causing dead neurons and preventing convergence to a global optimum. This issue is referred to as the vanishing gradient problem. To address these issues and enhance the training of deep networks, ResNet (Residual Networks) was introduced, incorporating residual blocks that allow for the learning of identity mappings and alleviate the degradation problem [122]. Residual blocks help in maintaining a stable gradient flow through the network, mitigating the risks of both exploding and vanishing gradients.

Figure 3.6 delineates the architecture of the residual Residual CNN employed in this research, comprising eight layers. Notably, all convolutional layers in the model adhere to uniform hyper-parameters: a kernel size of 3×3 and a total of 128 kernels. Given that the original SAR images undergo preprocessing via 4×4 average pooling to bolster computational efficiency, a stride size of 1 is chosen. This choice serves to attenuate the issue of mixed pixels and simultaneously augment the spatial resolution of the feature maps.

Subsequent to each convolution operation, Batch Normalization (BN) [123] is applied at every layer to normalize the output, thereby enhancing training efficiency. The activation function selected for IceNet is a variant of the Rectified Linear Unit (ReLU) [124], known as the Leaky Rectified Linear Unit (Leaky ReLU) [125]. Leaky ReLU diverges from the popular ReLU by allowing small negative values when the input is less than zero. This characteristic circumvents the "dying ReLU" issue, which manifests when a large gradient flows through the neurons, causing ReLU to invariably output zero, making a recovery from this state improbable. The leaky ReLU in IceNet is configured with a hyperparameter,



(a) (b)

Figure 3.6: Architecture of the labeling module in IceNet. (**a**) A residual block. (**b**) CNN architecture.

$\alpha = 0.2$, as per the empirical results presented in the study [125].

Three residual blocks follow the initial convolutional layer. As the main characteristic of residual CNN models, the identity mapping added by shortcuts in the residual block is depicted in Figure 3.6(a). Let x and y represent the input and output of the block. The residual block is defined as follows:

$$\begin{aligned} &F_i(x) = W_i x + b_i \\ &y = F_2(F_1(x)) + W_s x \end{aligned} \tag{3.3}$$

where $F_i(x)$ $W_i x$ and b_i are the convolution operation, learned weights and bias of ith layer. The block has two convolutional layers, and each layer has 128 kernels with the size of 3×3. The block requires the channel number of input x to be equal to that of $F_2(F_1(x))$, so they can pass the additive layer and send it to the activation function. If the dimensions are not the same, a 1 convolutional layer with the weight of W_s is added on the shortcut to change the dimension of x. The weights of each layer are initialized using the method proposed by He et al. [126]. Adam optimizer [127] is adopted for updating weights. The learning rate, the weight decay, and betas are set as 0.0001, 0, and [0.9, 0.999], respectively.

We select the cross-entropy cost function for the loss function for the model, which is described as follows:

$$loss = -\sum_{c}^{M} q_c log(p_c) \tag{3.4}$$

where M is the number of the classes, q_c is the expected output, p_c is the predict output of a softmax layer. Since this study focuses on classifying sea ice and open water, the loss function can be simplified into (3.5):

$$loss = -[q_c log(p_c) + (1 - q_c) log(1 - p_c)] \tag{3.5}$$

3.3.3 Regional Pooling Layer

When comparing the segmentation results with the corresponding CNN classification results, they deliver distinct outcomes. While the boundaries between ice and water are well-preserved in the former, the latter yields pixel-level labels. Leveraging the IRGS segmentation technique ensures that each region in the result image is treated as a homogeneous unit containing only a single class. Yet, the unsupervised nature of IRGS segmentation means these regions are only assigned with arbitrary classes.

To assign labels to these homogeneous regions based on CNN results, we propose incorporating a regional pooling layer. The first step involves defining the energy functions, E_w and E_i, of the homogeneous region, assuming that either water or ice is assigned to it, as follows:

$$E_w = -\sum_{a \in R_w} log(p_w^a) - \sum_{b \in R_i} log(1 - p_i^b) \tag{3.6}$$

$$E_i = -\sum_{a \in R_i} log(p_i^a) - \sum_{b \in R_w} log(1 - p_w^b) \tag{3.7}$$

Here, R_w and R_i denote the sets of pixels with hard labels of water and ice, respectively, as determined by the residual CNN in the region R–a homogeneous region in the segmentation result. a and b represent the elements in sets R_w and R_i, while p_w and p_i represent the probability values of being water or ice, respectively.

E_w and E_i are designed to use information content from information theory [128] to depict the energy of the region to which the label of water or ice is assigned. E_w represents how much energy the region would have if a water label is assigned to it, while E_i represents how much energy the region would have if an ice label is assigned to it. Thus, the labeling process is simplified to the determination of which label gives lower energy.

The energy terms E_w and E_i are designed to utilize the concept of information content from information theory[128], to characterize the energy of a region corresponding to water or ice labels, respectively. Specifically, E_w quantifies the energy of a region presuming a 'water' classification, whereas E_i quantifies the energy assuming an 'ice' classification. Thus, the labeling process is reduced to a comparison of energies, selecting the label that minimizes the energy for the region in question.

Subsequently, an energy term E_{iw} is computed to determine the correct label for R:

$$\begin{aligned} E_{iw} &= E_i - E_w \\ &= \sum_{a \in R_w} log(p_w^a) + \sum_{b \in R_i} log(1 - p_i^b) - \sum_{a \in R_i} log(p_i^a) - \sum_{b \in R_w} log(1 - p_w^b) \\ &= \sum_{c \in R} log(p_w^c) - \sum_{c \in R} log(p_i^c) \\ &= \sum_{c \in R} log(\frac{p_w^c}{p_i^c}) \end{aligned} \tag{3.8}$$

Where $c \in R = (R_w \cup R_i)$

This energy function is incorporated into a regional pooling layer. The label of R, represented as Y, is determined by E_{iw}. If $E_{iw} < 0$, $Y = ice$, otherwise $Y = water$, and every pixel in the region is then assigned the label Y. This process allows for an intricate integration of segmentation and classification results, thereby potentially enhancing the accuracy of sea ice classification in SAR images.

The introduction of the regional pooling layer draws intuitive inspiration from the pooling mechanisms prevalent in deep learning, as well as the processes used by CIS for producing sea ice charts. The segmentation output, generated through the application of IRGS, ensures that each region, or superpixel, is homogeneous in character. Subsequently, the labeling process, carried out by the residual CNN and regional pooling, parallels the method by which ice analysts annotate polygons on ice charts. A notable distinction, however, is that each superpixel requires labeling with only one class, a simplification made possible by its inherent homogeneity.

3.3.4 Comparative Methods

To further evaluate the performance of the proposed model for distinguishing sea ice and open water, we employ two benchmark methods widely used in machine learning studies and two referenced methods specially designed for sea ice classification from published papers for comparison. The two benchmark methods are residual CNN, which is also an essential component of IceNet, and random forest.

Referenced method 1: Leigh et al. [50] designed and implemented the SVM-IRGS for sea ice–water classification. SVM-IRGS is based on an MRF and adopts pixel labels predicted by SVM to modify the unary potential. A cross-validated feature selection is applied to the original feature set that includes GLCM features, backscattering intensities, and local averages and maximums to reduce over-fitting and improve computational efficiency. The 28 selected features are applied to train the SVM.

Referenced method 2: Hoekstra et al. [129] proposed the IRGS-RF model for distinguishing lake ice and water in SAR imagery. The model first oversegments the scene into homogeneous regions. Then, an RF classifier is used to assign labels to each region. The RF is trained using 162 GLCM features and ten backscattering features. This model is very similar to the method proposed in Chapter 2.

Recent studies have focused on distinguishing sea ice from open water, with many adopting deep learning frameworks. For instance, a dual-attention U-net model, referred to as DAU-Net, has shown promise in separating sea ice from open water due to its ability to capture semantic information within image blocks [114]. DAU-Net is anticipated to perform comparably to IceNet in terms of quantitative measures. However, the input block size for DAU-Net is fixed at 256 , which constrains the receptive field to a narrow scope. Consequently, this may result in misclassifications along the borders of image blocks when assembling the full scene. Additionally, Ma et al. [130] enhanced classification accuracy by integrating a CNN with CRF, yet this approach models the CRF over the entire scene, rendering it sensitive to variations in the incidence angle. This sensitivity could lead to noise-like misclassifications throughout the generated sea ice maps. Owing to these limitations, DAU-Net and the CNN-CRF model are not selected for comparative analysis in this study.

3.4 Results and Discussion

In order to evaluate the performance of IceNet presented in this study, we compared IceNet with RF, residual CNN, SVM-IRGS, and IRGS-R. The experiments were assessed with accuracy for each scene in the dataset. The accuracy is defined as follows.

$$\text{accuracy} = \frac{TP + FN}{N} \times 100\% \tag{3.9}$$

where TP and FN are the numbers of true-positive and false-negative samples, and N is the total number of samples. The overall accuracy for each method is calculated using all samples in the dataset.

The experiments were run on a computer with the following configuration: INTEL Core i5-6600K CPU, 32-GB RAM, NVIDIA GeForce GTX 1080 GPU, and Windows 10 operating system. The average training time is 4–5 hours under such configuration. It takes 15 minutes to produce a sea ice map for each SAR scene, which is 2–3 min for segmenting SAR imagery, 10 for predicting pixel-level labels using residual CNN, and 2–3 min for regional pooling and producing the final sea ice map.

3.4.1 Classification Accuracy

residual CNN is selected as a benchmark method in this article, and IceNet also relies on the spatial information learned by residual CNN; therefore, optimizing the residual CNN model is essential. We investigate different setups for the residual CNN to achieve the best classification accuracy. Overall accuracy is achieved based on the LOO approach, while other accuracies without specific indication are obtained using the whole dataset.

The backscattering observed in SAR imagery is predominantly influenced by the dielectric properties and surface roughness of sea ice/water, as well as the underlying backscattering mechanisms. While HH polarization tends to offer richer information for distinguishing between sea ice types compared to HV polarization, HV polarization images typically exhibit lower intensity values. This results in reduced absolute variations in the statistical properties of HV polarization, attributable to the multiplicative nature of speckle noise, as opposed to HH polarization. This characteristic of HV polarization enhances its capability to differentiate various sea ice types [52]. In light of this, we exploit how polarization mechanisms impact the performance of deep learning models.

To this end, we have leveraged different combinations of polarized images as inputs for the residual CNN model. Specifically, image patches extracted from HH, HV, and combined HH/HV polarizations have been utilized to assess the classifier's performance. Table 3.3 presents the validation accuracy achieved using an image patch size of 25×25 pixels. The results demonstrate that the HV polarization scene outperforms the HH polarization scene in terms of classification accuracy when a single polarized image is employed as input. Notably, the combination of HH and HV polarizations for image patches culminates in the highest overall accuracy, recorded at 98.65%. In light of these findings, subsequent experiments will incorporate both HH and HV scenes as inputs for the residual CNN and IceNet models.

Table 3.3: Overall accuracy of residual CNN using different polarization combinations, HH, HV, and HH/HV. The patch size is 25×25.

Input Channel	**HH**	**HV**	**HH/HV**
Validation Accuracy	93.82%	95.93%	**98.65%**

After the architecture of residual CNN has been decided, the receptive field of the model is determined by the input patch size. A Large patch size can capture more spatial information that contributes to feature maps than a small one, but it may contain mixed ice types and water that confuse the classifier. We consider using different patch sizes of 5×5, 13×13, 25×25, 39×39, and 51×51 pixels to evaluate the performance of residual CNN. Figure 3.7 shows the validation accuracy using different patch sizes. In general, the accuracy increases with a larger patch size; however, the benefit of using the patch size of 51×51 is limited. First, a 51×51 patch covers an area of 10,200 by 10,200 m. Such extensive coverage may capture both sea ice and water that would be expected to cause classification errors at the boundaries between different ice types and open water. Moreover, a large patch size requires more computation to process. For example, the 39×39

patch only improves the accuracy by 0.42% compared with the 25×25 patch with doubled running time; therefore, after balancing the trade-off between patch size and classification performance, the smaller patches (5×5, 13×13, and 25×25) are selected for further comparison.

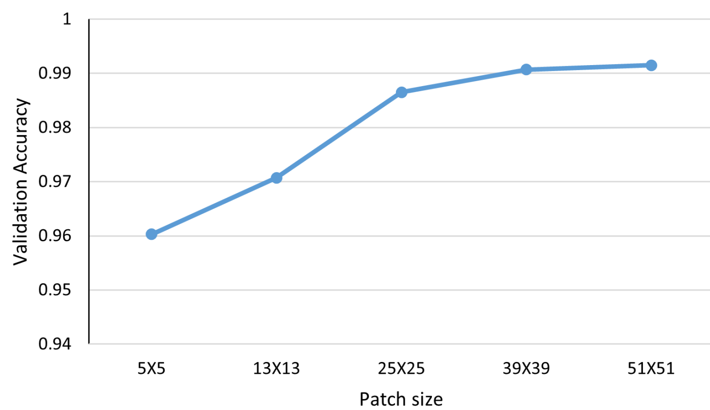


Figure 3.7: Validation accuracy using patch size of 5×5, 13×13, 25×25, 39×39, and 51×51.

The proposed IceNet and comparative methods can be sorted into two categories: single model and combined model. residual CNN and RF are pixel-wise classifiers that belong to the single-model category. In contrast, IceNet, SVM-IRGS, and IRGS-RF are constructed by using both segmentation and labeling and are defined as combined models. Table 3.4 shows classification accuracy for each scene and overall accuracy using the aforementioned methods. The highest classification accuracy obtained for each scene is highlighted. If several methods achieve the same highest accuracy, no one will be highlighted for this scene. In the single-model category, RF demonstrates a remarkable performance, achieving an overall classification accuracy of 96.19%. This result notably surpasses that of the residual CNN when a small patch size of 5×5 is utilized. Nevertheless, as the patch size is incrementally increased, there is a shift in performance. The classification performance of the residual CNN begins to exceed that of RF. This enhancement in performance can be attributed to the expansion of the receptive field, which allows the residual CNN to capture and integrate information from a larger spatial context, thereby significantly improving its classification accuracy.

For more complex models, the classification accuracy of SVM-IRGS and IRGS-RF achieved are consistent with the experimental results reported in the original studies [50, 129]. IceNet improves the overall accuracy by 1.86%, 0.39%, and 0.02% compared with residual CNN using patch size of 5 × 5, 13 × 13, and 25 × 25. Though the improvement accomplished by IceNet is not impressive in terms of overall accuracy, the classification results predicted by IceNet are the runner-up for classification accuracy consistency among all methods in this study. Figure 3.8 shows the box and whisker plot of the distribution of classification accuracy obtained by these models for all 21 scenes. The boxes in the plot represent the interquartile range, which is the range between the first quartile and the third quartile. Residual CNN (5 × 5 patch), RF IceNet (5 × 5 patch), SVM-IRGS, and IRGS-RF struggle with several challenging scenes, while IceNet achieves the best classification accuracy with minimum variance.

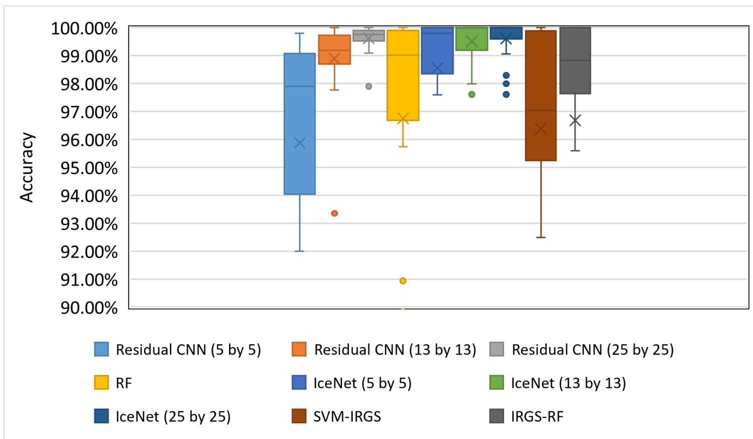


Figure 3.8: The distribution of classification accuracy achieved by IceNet and comparative methods for all 21 scenes shown in the box plot. Outliers are represented by dots, and "X" is the mean value. Several outliers in residual CNN (5×5), RF IceNet (5×5), SVM-IRGS, and IRGS-RF are ignored for the layout of the plot.

3.4.2 Ice–Water Maps

Although the differences in numerical accuracy between IceNet and the comparative methods are marginal, the robust performance of the proposed method could not be entirely depicted using the numerical result, which is computed based on limited, high-confidence samples. To complement the classification accuracies (based on a small set of samples), it is prudent to assess the results on the entire scene visually.

The scene results acquired on 3 October 2010, displayed in Figure 3.9, deliver a visual example of full scene classification that could be deployed for operational use. The HV

Table 3.4: Classification accuracy for 21 scenes using residual CNN (patch size 5×5, 13×13, and 25×25), RF, IceNet (patch size 5×5, 13×13, and 25×25), SVM-IRGS, IRGS-RF.

Scene ID	Single model				Combined model				
	Residual CNN			RF	IceNet			SVM-IRGS	IRGS-RF
	5×5 patch	13×13 patch	25×25 patch		5×5 patch	13×13 patch	25×25 patch		
20100418_163315	98.28%	99.87%	99.84%	98.86%	100.00%	100.00%	100.00%	100.00%	100.00%
20100426_040439	99.41%	99.80%	99.89%	98.93%	100.00%	100.00%	100.00%	100.00%	100.00%
20100510_035620	98.46%	99.65%	99.94%	98.91%	100.00%	100.00%	100.00%	99.49%	100.00%
20100524_034756	97.90%	98.93%	**99.08%**	97.06%	98.25%	98.29%	98.29%	98.52%	98.30%
20100605_163323	98.65%	98.94%	99.78%	97.61%	98.45%	98.45%	**99.99%**	99.32%	97.68%
20100623_041255	99.23%	99.54%	99.66%	97.84%	99.79%	99.99%	**100.00%**	98.31%	99.80%
20100629_163326	98.90%	99.26%	99.75%	90.93%	99.52%	99.61%	99.48%	96.10%	**99.93%**
20100712_031834	97.61%	98.43%	99.18%	98.65%	100.00%	100.00%	99.71%	95.74%	98.21%
20100721_173208	99.79%	100.00%	100.00%	98.41%	97.99%	97.99%	97.99%	95.25%	97.99%
20100730_162908	91.99%	97.76%	99.58%	72.82%	98.91%	99.88%	**100.00%**	93.05%	73.34%
20100807_173610	86.44%	98.54%	99.69%	96.97%	100.00%	100.00%	100.00%	92.49%	100.00%
20100816_163329	78.08%	93.35%	97.90%	76.82%	82.02%	99.58%	**100.00%**	75.63%	77.47%
20100907_035614	99.56%	99.93%	99.97%	99.04%	100.00%	100.00%	100.00%	99.78%	100.00%
20100909_163321	99.61%	99.97%	**100.00%**	97.89%	97.59%	97.61%	97.61%	94.70%	97.59%
20101003_163324	92.87%	99.10%	99.63%	94.29%	99.96%	100.00%	100.00%	96.65%	98.37%
20101021_041325	95.21%	98.86%	**99.77%**	95.07%	99.48%	99.71%	99.71%	97.04%	98.83%
20101027_025726	98.48%	99.30%	99.45%	97.82%	99.71%	99.75%	**99.76%**	96.70%	99.78%
20101114_041304	92.24%	97.79%	**99.45%**	92.74%	98.15%	98.80%	99.05%	95.24%	95.59%
20101120_163324	97.49%	99.41%	99.91%	96.21%	100.00%	100.00%	100.00%	100.00%	97.56%
20101206_015139	97.47%	99.18%	99.60%	97.01%	99.84%	100.00%	99.99%	100.00%	99.99%
20101214_025725	95.63%	99.07%	99.77%	96.30%	100.00%	100.00%	100.00%	100.00%	99.77%
Overall	97.03%	99.07%	99.65%	96.19%	98.89%	99.46%	**99.67%**	96.28%	97.75%

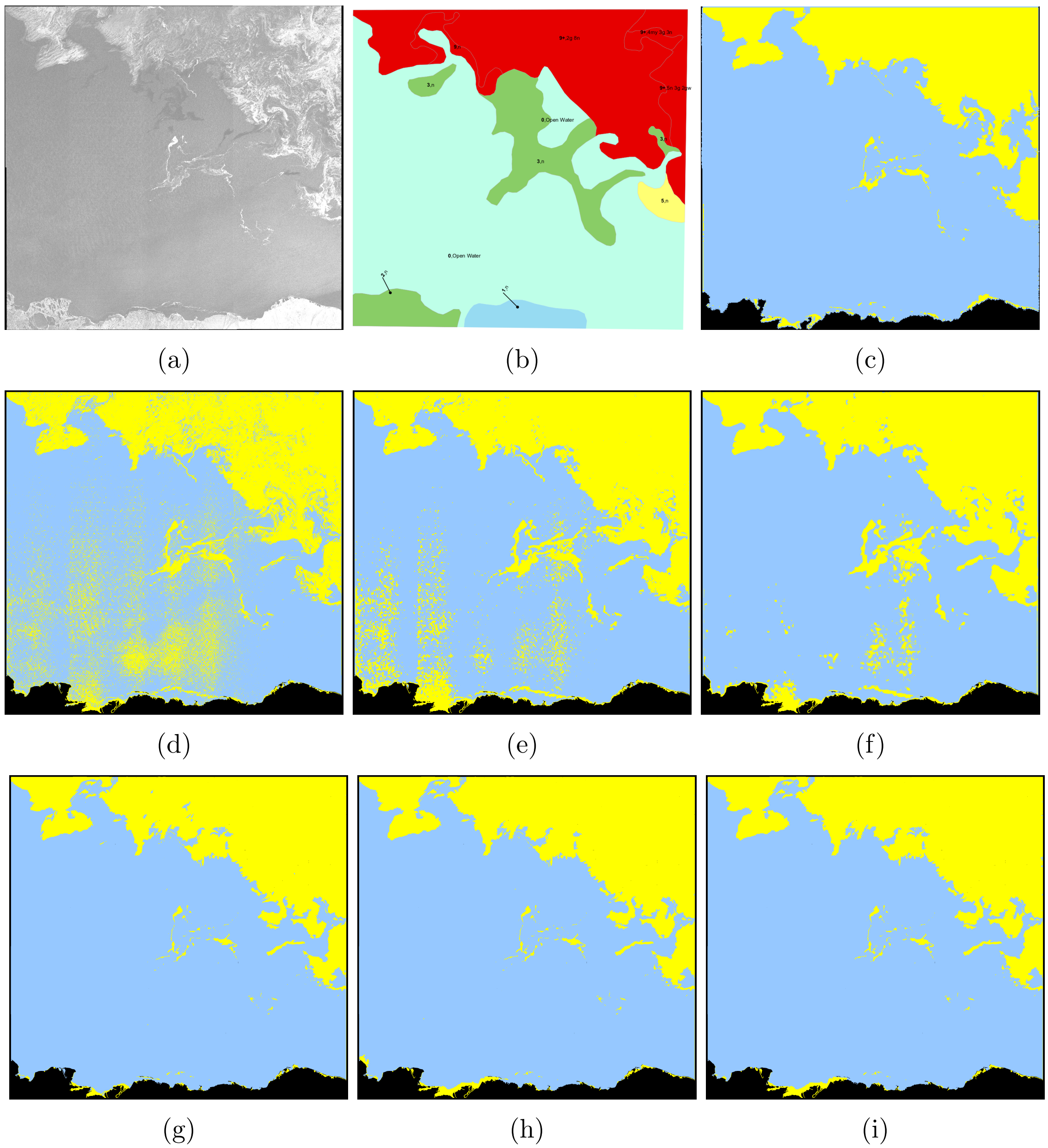


Figure 3.9: Classification results of 3 October 2010. Water (blue), ice (yellow), and land mask (black). (**a**) HH polarization. (**b**) Ice chart. (**c**) SVM-IRGS: 96.65%. (**d**) residual CNN with the patch size 5: 92.87%. (**e**) Residual CNN with the patch size 13: 99.10%. (**f**) Residual CNN with the patch size 25: 99.63%. (**g**) IceNet with the patch size 5: 99.96%. (**h**) IceNet with the patch size 13: 100.00%. (**i**) IceNet with the patch size 25: 100.00%.

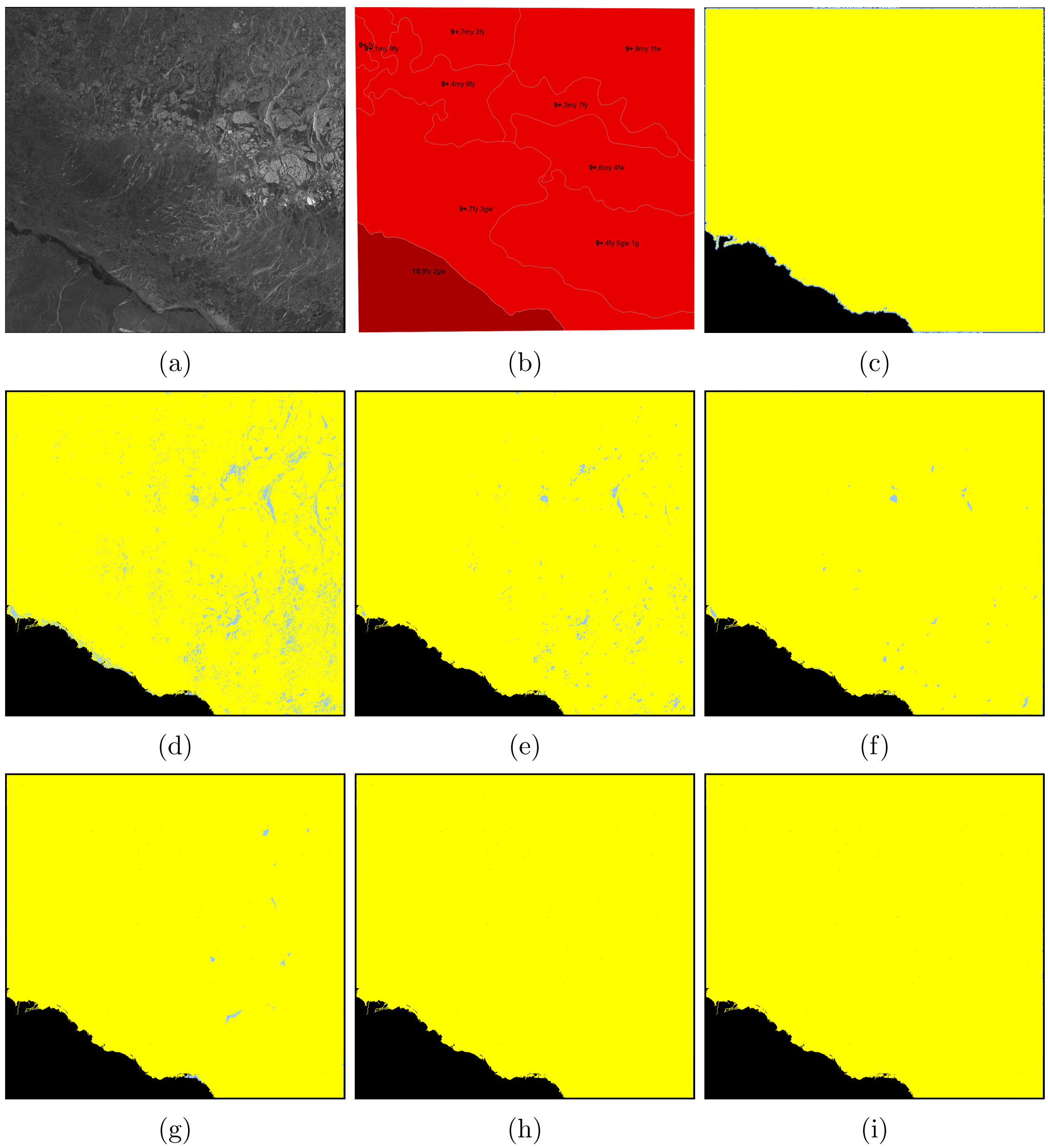


Figure 3.10: Classification results of 14 December 2010. Water (blue), ice (yellow), land mask (black). (**a**) HH polarization. (**b**) Ice chart. (**c**) SVM-IRGS: 100.00%. (**d**) Residual CNN with the patch size 5: 95.63%. (**e**) Residual CNN with the patch size 13: 99.07%. (**f**) Residual CNN with the patch size 25: 99.77%. (**g**) IceNet with the patch size 5: 100.00%. (**h**) IceNet with the patch size 13: 100.00%. (**i**) IceNet with the patch size 25: 100.00%.

image is contaminated by noticeable inter-scan banding, a common artifact presented in SAR imagery captured under ScanSAR mode [12]. Due to the scanning mechanism, the antenna of the SAR system transmits and receives multiple beams to obtain a wide swath under ScanSAR mode; however, the backscattering of these beams is different near the borders between scans due to temporal variants of the antenna pattern. These variants appear on the SAR imagery as inter-scan banding artifacts and cause the inconsistency of backscattering across the whole scene. Figure 3.9(d) to 3.9(f) illustrate that all three residual CNNs have reduced classification accuracy derived from the inter-scan banding artifact. The residual CNNs are confused by the artifacts and misclassified sea ice as open water. With increasing patch size, the negative impact associated with the banding artifact is visually reduced. SVM-IRGS overcomes this problem by combining segmentation with labeling; however, there are still some water errors presented around the ice–water boundary.

The improvement achieved by IceNet is significant. Classification errors associated with the banding artifact are mitigated, and the ice–water boundaries retain the naturally occurring details. Since a 5×5 patch cannot capture sufficient spatial extent, many water classification errors appear in the top-middle of the scene (Figure 3.9(d)). IceNet corrects these water errors by introducing contextual information learned from IRGS segmentation. The water errors presented in the results of SVM-IRGS are also refined in those of IceNet. The highest classification accuracy for this scene is 100.00% achieved by IceNet using 13×13 and 25×25 patch.

An example from the ice freeze-up season acquired on 14 December 2020 is shown in Figure 3.10, where only ice and land appear in the scene. Although different ice types, young (grey and grey-white) ice, first-year ice, and multi-year, are presented, this is the least challenging scene in the dataset. Since the scene was captured in December, the ice condition was not stable, and ridges and fissures appeared across the whole scene. Residual CNNs fail to classify the leads containing newly formed sea ice among multi-year ice because the new ice formations have a similar texture to water, and their backscatter is lower compared to the surrounding ice. Residual CNN with a patch size of 5×5 obtains the lowest accuracy due to the limited spatial extent. As the patch increases, the classification errors are rectified; however, there are still water errors residual in the ice map produced by residual CNN with a patch size of 25×25. On the contrary, the proposed IceNet eliminates all water errors successfully using patch sizes of 13×13 and 25×25. Similar classification results are also obtained in some other scenes, e.g., scene ID 20101120_163324, 20101206_015139.

Classifying sea ice during the melt seasons is usually the most challenging time in the life span of sea ice. Figure 3.11 depicts the classification results of the scene acquired on 30 July 2020. Since the patch size of 25 × 25 achieves the best results in both visual interpretation and overall accuracy, only residual CNN and IceNet using the patch size of 25 × 25 are presented in the figure. Both sea ice and open water show significantly inconsistent appearances across the whole scene, posing challenges to classification with high accuracy. The texture of sea ice in the scene is degraded by melting and looks similar to open water; therefore, the multi-year ice presented in the upper left corner is misclassified as open water by IRGS-RF. Again, The IceNet is robust to these intra- and inter-class variances and achieves the highest classification accuracy of 100%.

Boundary preservation is also critical for ice–water classification. Figure 3.12 displays a region of interest extracted from Figure 3.11, and the boundaries between sea ice and open water are highlighted in red. Although residual CNN with a patch size of 5 × 5 detects the outlines of big ice floes, numerous water errors appear inside these floes, and noise-like ice errors are also present in open water. As the patch size increases, these classification errors are mitigated, but the boundaries between small ice floes and water are ignored by residual CNN. In contrast, the proposed IceNet shows consistent performance in detecting ice–water boundaries through different patch sizes.

Classifying sea ice during the melt seasons is usually the most challenging time in the life span of sea ice. Figure 3.11 depicts the classification results of the scene acquired on 30 July 2020. Since the patch size of 25 × 25 achieves the best results in both visual interpretation and overall accuracy, only residual CNN and IceNet using the patch size of 25 × 25 are presented in the figure. Both sea ice and open water show significantly inconsistent appearances across the whole scene, posing challenges to classification with high accuracy. The texture of sea ice in the scene is degraded by melting and looks similar to open water; therefore, the multi-year ice presented in the upper left corner is misclassified as open water by IRGS-RF. Again, The IceNet is robust to these intra- and inter-class variances and achieves the highest classification accuracy of 100%.

Boundary preservation is also critical for ice–water classification. Figure 3.12 displays a region of interest extracted from Figure 3.11, and the boundaries between sea ice and open water are highlighted in red. Although residual CNN with a patch size of 5 × 5 detects the outlines of big ice floes, numerous water errors appear inside these floes, and noise-like ice errors are also present in open water. As the patch size increases, these classification errors are mitigated, but the boundaries between small ice floes and water are ignored by residual CNN. In contrast, the proposed IceNet shows consistent performance in detecting ice–water boundaries through different patch sizes.

In general, although ice errors caused by the banding artifact and water errors caused by the complex backscattering are observed in the sea ice maps produced by residual

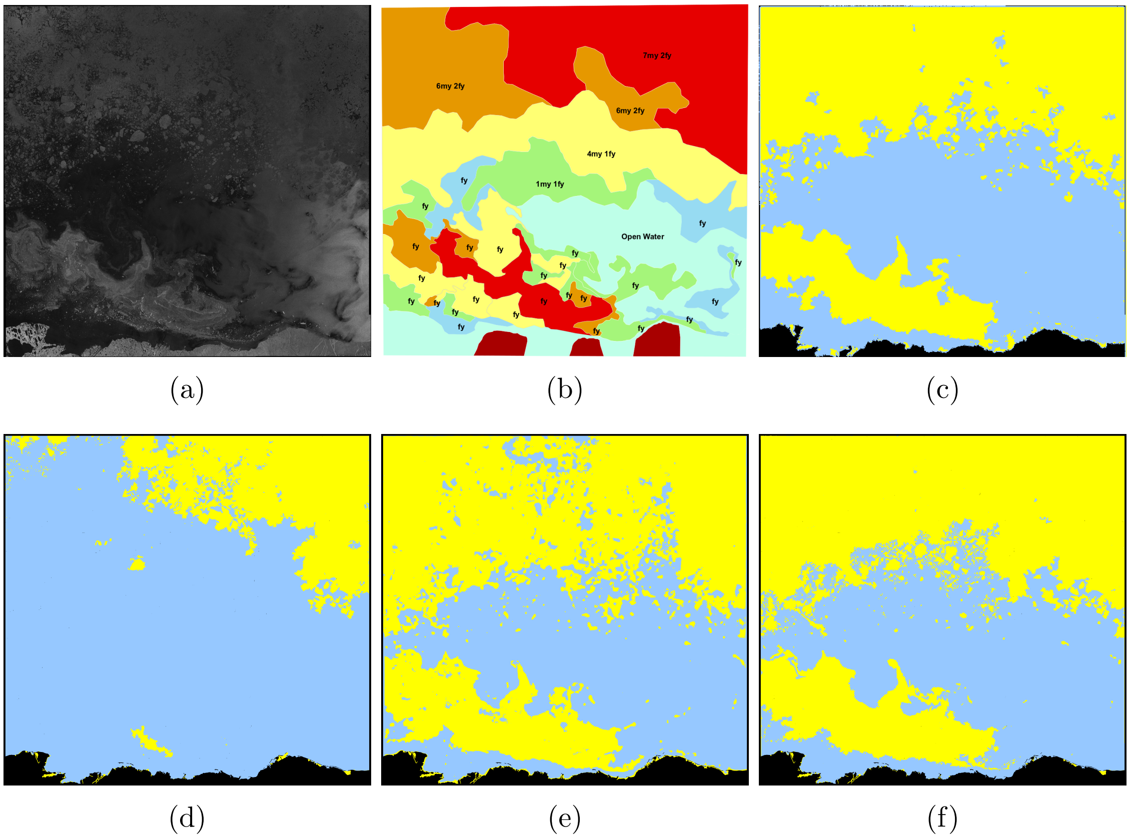


Figure 3.11: Classification results of 30 July 2010 (scene ID 20100730_162908. Testing). Water (blue), ice (yellow), land mask (black). (**a**) HH polarization. (**b**) Ice chart. (**c**) SVM-IRGS: 93.05%. (**d**) IRGS-RF: 73.34%. (**e**) Residual CNN with the patch size 25: 99.58%. (**f**) IceNet with the patch size 25: 100.00% .

CNNs, the overall accuracy still outperforms the methods for comparison. The proposed IceNet achieves the highest classification accuracy of 99.67%. Even though the numerical improvement is neglectable, the contribution of IceNet can be demonstrated in generating sea ice maps: the ice–water boundary is refined, and the ice and water errors are corrected.

3.5 Summary

The robust and automatic IceNet classification method is proposed in this paper to classify sea ice and open water. A regional pooling layer is applied to combine the unsupervised

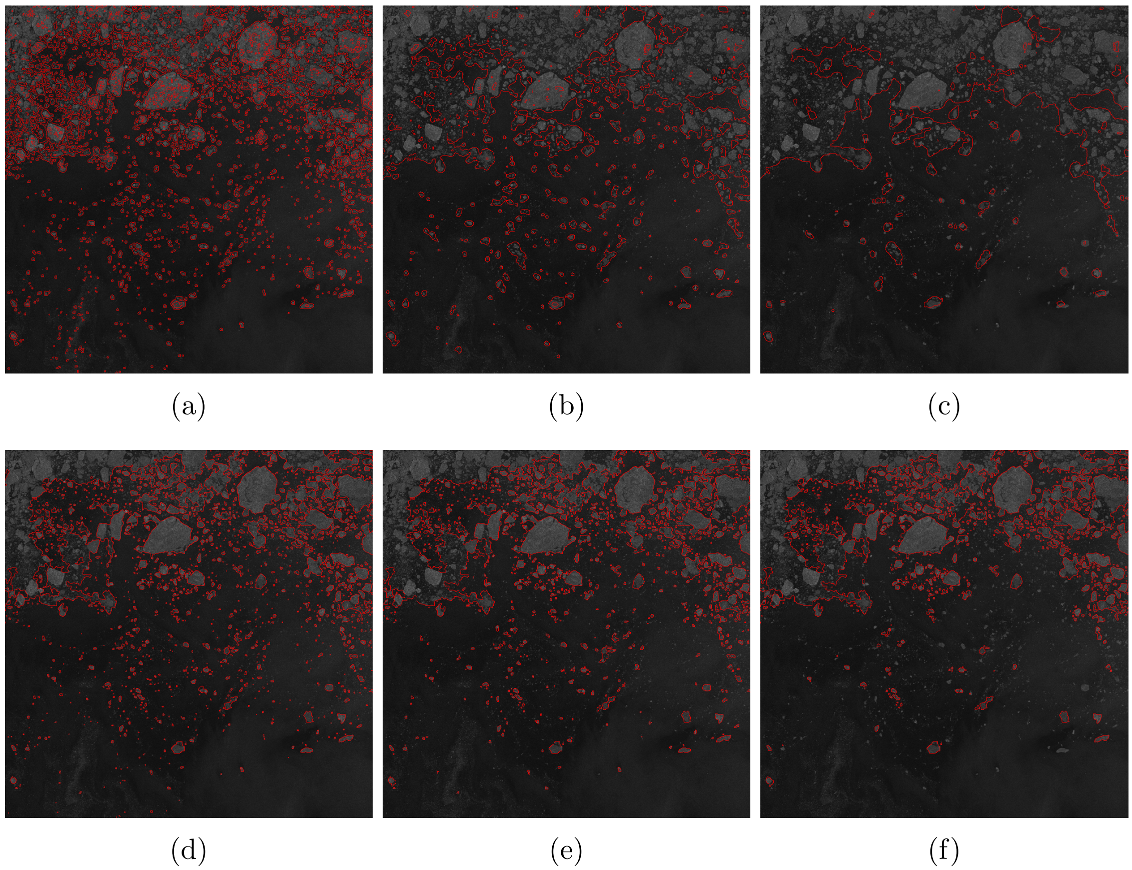


Figure 3.12: Highlighted ice–water boundary of 30 July 2010 (scene ID 20100730_162908). (**a**) Residual CNN with the patch size 5. (**b**) Residual CNN with the patch size 13. (**c**) Residual CNN with the patch size 25. (**d**) IceNet with the patch size 5. (**e**) IceNet with the patch size 13. (**f**) IceNet with the patch size 25.

IRGS segmentation and supervised pixel-wise residual CNN labeling, both of which are state-of-the-art methods in remote sensing. The performance of the IceNet is evaluated on 21 RADARSAT-2 scenes of the Beaufort Sea from 2010 with two benchmark methods and two referenced methods for comparison.

For single-model classifiers, the residual CNN with a patch size of 25×25 demonstrates superior performance, achieving an overall accuracy of 99.65%. These results suggest that, given ample labeled data, deep learning models generally surpass traditional RF classifiers. This finding also underscores the proficiency of spatial features learned by the residual

CNN in differentiating between sea ice and open water, as opposed to the handcrafted GLCM textural features, which necessitate expert knowledge for feature design. The self-taught feature learning by the residual CNN exhibits robustness in complex scenes, with its classification accuracy outstripping that of combined-model approaches such as SVM-IRGS and IRGS-RF. Capitalizing on the residual CNN's formidable performance, IceNet further elevates the overall accuracy to 99.67% with a 25×25 patch, thereby affirming the significance of integrating contextual information to enhance classification precision.

Since the numerical accuracy is calculated using limited labeled samples and could not fully demonstrate the performance of the proposed method, the sea ice maps generated by IceNet are also assessed by visual analysis. According to visual inspection, the noise-like ice errors caused by inter-scan banding are suppressed, and the ice–water boundaries are refined with natural details. The water errors in residual CNN's results are also ameliorated.

In spite of the robust performance achieved by the proposed IceNet, some limitations exist. First, the IceNet is only tested for ice–water classification due to limited labels of sea ice types. It will be evaluated to distinguish different sea ice types in future work. Second, only local relation is considered for regional pooling. The global relation between regions should be taken into account for better classification performance.

Chapter 4

An Interactive Sea Ice Classification Method for SAR Imagery Based on Convolutional Neural Network and Graph Convolutional Network

Monitoring sea ice in the Arctic region is crucial for polar maritime activities. The Canadian Ice Service (CIS) wants to augment its manual interpretation with machine-learning-based approaches due to the increasing data volume received from newly-launched synthetic aperture radar (SAR) satellites. However, fully supervised machine learning models require large training datasets, which are usually limited in the sea ice classification field. To address this issue, a semi-supervised interactive system to classify sea ice in dual-pol RADARSAT-2 imagery using limited training samples is proposed in this chapter. First, the SAR image is oversegmented into homogeneous regions. Then, a graph is constructed based on the segmentation results, and the feature set of each node is extracted by a convolutional neural network. Finally, a graph convolutional network (GCN) is employed to classify the whole graph using limited labeled nodes automatically. The proposed method has been evaluated on a published dataset. Compared with referenced algorithms, this new method outperformed in both qualitative and quantitative aspects.

4.1 Introduction

Arctic sea ice plays a vital role in the global climate and local ecosystems. It reflects the incoming solar radiation away and prevents energy exchange between oceans and the atmosphere to cool the polar region. Local communities and marine mammals rely on sea ice for hunting, traveling, and other daily activities [131]. Due to global warming, the sea ice extent has declined rapidly, with a rate of 13% per decade [8]. The 15 lowest sea ice extent yearly minimums on record have all occurred in the past 15 years [132, 133]. Nevertheless, with less ice covering the Arctic Ocean, shipping routes that were once inaccessible or dangerous have become more viable, bringing remarkable economic benefits for ocean transportation [134]. The safety of Arctic shipping requires accurate monitoring of sea ice as well. Therefore, the continuous mapping of sea ice extent and how it changes over time has become a crucial research topic.

Among data collected from different space-borne remote sensors, synthetic aperture radar (SAR) imagery has been demonstrated to be reliable for sea ice remote sensing [135, 136]. As an active radar, SAR offers moderate spatial resolution and expansive coverage regardless of polar darkness and weather conditions. To date, national ice agencies, such as the Canadian Ice Service (CIS), the US National Ice Center (NIC), and the Greenland Ice Service affiliated with the Danish Meteorological Institute (DMI), rely mainly on SAR data to provide information on ice conditions to users in the form of ice charts. A manually drawn ice chart usually covers the ice concentration, stage of development, and form of ice for the matched SAR image. Data collected from sources (e.g., passive microwave data, environmental reanalysis data, and in-situ observations) are also used as references for producing ice charts. Although the quality of ice charts is well controlled by ice analysts, the labeling process is labor-intensive and time-consuming, and thus, the number of ice charts that can be produced from SAR images on a given day is limited [137]. Therefore, to produce more ice maps that cover a larger area with higher temporal resolution, it is desirable to have a process in place that can either fully or partially automate the analysis of SAR sea ice imagery [113].

As one of the most important tasks for sea ice mapping, sea ice classification consists of two steps: ice cover detection and ice typing [138]. Although numerous studies for automatic/semi-automatic sea ice cover detection (i.e., ice-water classification) have been presented with high reported classification rates, distinguishing different ice types is a more challenging step. One of the main challenges is the overlap of backscattering signatures of different sea ice types. For example, at the C-band, the HH-polarized microwave backscatter coefficient increases from grey to grey-white ice and then decreases as the ice grows [139]. This indicates that using only backscatter intensity is not sufficient to discriminate sea ice types. Moreover, the statistical non-stationarity introduced by the change in backscatter intensity as a function of the incidence angle causes backscatter variation of any particular sea ice type across the SAR scene [98]. During freeze-up and melt periods, classification becomes increasingly difficult due to wet snow lowering radar penetration depth, snow metamorphism, and increased ice dynamics [140].

To develop an automatic/partially automatic and robust sea ice classification system that overcomes the aforementioned challenges, researchers have been applying machine learning methods with features extracted from SAR imagery. Polarimetric and textural features derived from the gray-level co-occurrence matrix (GLCM) [24] have been used with classifiers such as Bayesian classifier [52], decision tree [35], support vector machine (SVM) [34, 7], conditional random field (CRF) [45, 141, 142, 48], and Markov random field (MRF) [49]. In recent years, deep learning has become popular in remote sensing from SAR imagery. Among different data-driven deep learning models, convolutional neural networks (CNNs) are widely adopted for sea ice classification [41, 111, 102, 10, 143, 40, 144] and sea ice concentration estimation [63, 145, 137, 146, 147]. The ability to learn robust features automatically from a large volume of training data makes the CNN-based model a more preferable choice for sea ice classification compared with traditional machine learning.

However, so far, none of the CNN-based methods for sea ice classification have been applied for operational sea ice mapping. One reason is that their classification is conducted on a pixel-wise level, which is inconsistent with the 'polygon'-based format of operational ice charts. Specifically, the ice analyst manually demarcates the full SAR scene into appropriate spatial regions called polygons. Then, the analyst interprets each polygon, assigning codes to define ice concentration and stages of development according to the sea ice nomenclature defined by the World Meteorological Organisation (WMO) [1]. Another reason is that due to a lack of sufficient ground-truth pixel-based samples, those CNN-based methods are trained using limited SAR samples. Whether they can produce reliable predictions over the data collected from different times and locations is uncertain. In contrast, an ice chart is required for review by multiple experts before being distributed to the public, and an amended or corrected version will be released if necessary [6], which further ensures its reliability.

In contrast to the pixel-level automatic classification that specifies which pixel belongs to which class without quality control, it is worthwhile to pursue a method that provides regional-level (i.e., polygon-based) labels that specify which type of ice is contained in a region. With polygon-based information from ice charts, the region labels are more robust to identification errors and easier to acquire [148]. Besides, to tackle the issue of limited training data, semi-supervised learning has been introduced [149]. Various publications demonstrate that semi-supervised methods, including self-training [150], semi-supervised

SVM [151], shared subspace learning [152], and graph-based neural networks [153, 154], achieve robust performance when dealing with limited labeled training data. Li *et al.* [142] presented an ice-water classification method called ST-IRGS, which integrates semantic segmentation, global merging, and self-training. The algorithm outperformed the Gaussian maximum likelihood classier and Gauss-Markov random field on a dual-pol RADARSAT-2 dataset with scarce training samples. Khaleghian *et al.* [11] reported a teacher-student-based semi-supervised deep learning method to discriminate sea ice types. The proposed method learned sea ice characteristics from limited labeled samples and massive unlabeled samples.

Graph-based neural networks, especially graph convolutional networks (GCNs) [155], stand out from the rest of the semi-supervised methods for classification in remote sensing imagery. Specifically, pixel-based methods cannot capture the inherent geometry and distinct structure in the remote sensing data space. The vertices and edges in a graph naturally represent the topological relationships in the remote sensing imagery, and GCNs can extract non-Euclidean features from each vertex. Moreover, the graph structure in a GCN allows for significant computational cost reduction compared with pixel-based methods. Zhang *et al.* [156] developed a GCN-based model named SPGCN for hyperspectral image classification. A spatial pooling layer was introduced to the model to reduce the patch size and graph size after each convolutional layer. The model's performance was evaluated on three hyperspectral datasets, and the results illustrated that SPGCN achieved competitive accuracy compared with CNN-based models with less runtime. Wang *et al.* [157] applied broad learning as a fully connected layer to GCN and used an intra-class divergence matrix and an inter-class divergence matrix to train it. The proposed model considered both the inter-class and intra-class spacing of sample features and improved the classification accuracy for hyperspectral images compared with that of a classic GCN.

While the methodologies presented in Chapters 2 and 3 yield encouraging results in sea ice classification, they are subject to certain limitations. Firstly, the integration of spatial and contextual features is somewhat akin to a post-processing step. The labeling mechanisms within these methods depend solely on the segmentation outcomes, restricting the potential synergistic interplay between spatial and contextual information that could enhance classification accuracy. Secondly, the scope of contextual features in these approaches is confined to local regions, leaving the influence of global contextual information on classification performance unexplored.

To address these challenges and establish a more effective framework for sea ice classification, this thesis introduces an interactive sea ice classification method that discerns between water and three types of ice. This method, termed IceGCN, integrates a CNN with a GCN. Unlike conventional handcrafted features, such as those based on GLCM, the

CNN-derived features elevate classification precision while accelerating the feature extraction process. Furthermore, the IRGS unsupervised segmentation algorithm is employed to retain the delineation of boundaries between different ice types. In addition, the GCN component of IceGCN is designed to learn contextual features adeptly on a global scale, surpassing the regional limitations of previous methods. Employing a semi-supervised GCN allows for the fusion of spatial and contextual features, generating accurate sea ice classification maps with a constrained number of labeled samples, thus aligning with operational needs. The efficacy of the proposed IceGCN approach is rigorously assessed in comparison to established benchmarks using a dual-polarization RADARSAT-2 dataset.

Therefore, to address these challenges and provide a more robust and efficient pipeline for sea ice monitoring, we propose an interactive sea ice classification method to identify water and three ice types. The method integrates a CNN into a GCN and is named IceGCN. Compared with traditional handcrafted features, e.g., GLCM features, applying the features extracted by CNN improves the classification accuracy and reduces the processing time of feature extraction. Moreover, boundaries between ice and different ice types are preserved by introducing the Iterative Region Growing with Semantics (IRGS) unsupervised segmentation algorithm. Unlike the supervised models used in other studies, a semi-supervised GCN is employed to combine the spatial context features and produce reliable sea ice classification maps using limited labeled samples for operational purposes. The performance of the proposed method is evaluated and compared with benchmark methods on a dual-pol RADARSAT-2 dataset.

Following Section 4.1, Section 4.2 describes the SAR dataset used in this research and presents the proposed sea ice classification system in detail. The experimental results are discussed in Section 4.3. In the end, the conclusion is summarized in Section 4.4.

4.2 Methodology

In this section, we present a semi-supervised method that seamlessly combines local spatial characteristics extracted through CNN with global spatial features derived from GCN for the purpose of classifying sea ice. The workflow of our proposed method, named IceGCN, is illustrated in Figure 4.1. First, the input HH/HV scenes are oversegmented into small homogeneous regions (superpixels) by employing the IRGS algorithm [70]. Subsequently, a graph can be constructed upon these superpixels, wherein the weights between adjacent superpixels are determined by edge strength. Following this step, a pre-trained CNN extracts pixel-level features that are then aggregated by pooling to generate an array of feature vectors, each representing the essence of an individual superpixel. Later, two

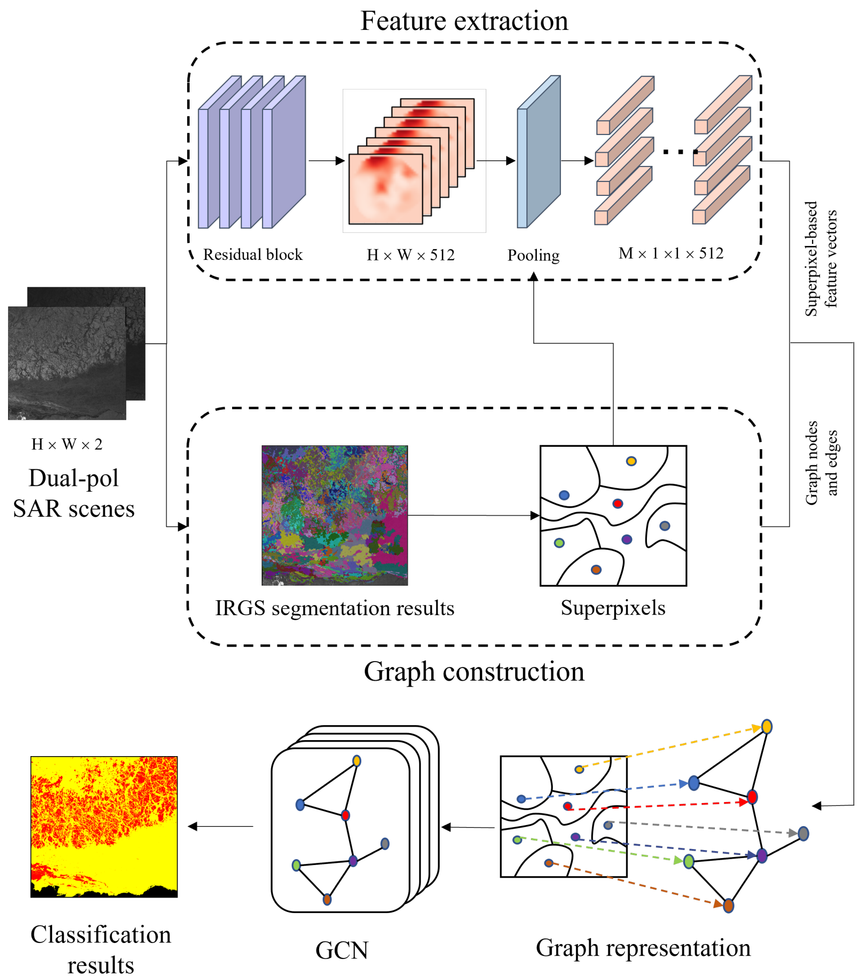


Figure 4.1: Workflow diagram of the proposed IceGCN.

graph convolutional layers are deployed to learn the spatial and contextual relationships between superpixels, utilizing a limited amount of labeled data. Finally, a softmax layer is introduced to ingest outputs of the graph convolutional layers and assign the sea-ice label for each superpixel. Unlike the two-stage frameworks previously posited in this thesis, which distinctly segment and then label the SAR imagery, IceGCN innovatively merges segmentation, and labeling into a singular, cohesive stage.

4.2.1 Superpixel Generation

SAR imagery that is commonly utilized for sea ice monitoring often covers broad spatial expanses, with pixel dimensions reaching up to 10,000 by 10,000. Building a graph on such an enormous number of pixels becomes impractical. Thus, we divide the SAR imagery into smaller regions known as superpixels, composed of pixels sharing highly similar characteristics. By constructing a graph on these superpixels, we significantly decrease the required memory space and computational power. The IRGS method, a Markov random field (MRF)-based algorithm specifically designed to provide reliable segmentation in SAR imagery, is utilized to generate superpixels in this study. The segmentation is performed by minimizing an energy function that blends the Gaussian mixture distribution and edge strength. The energy function is defined as follows [70]:

$$E = \sum_{i \in R} \upsilon_G(x_i) + \sum_{\langle i,j \rangle \xi} \upsilon_E(x_i, x_j) \tag{4.1}$$

where $\upsilon_G(\cdot)$ depicts the Gaussian statistics for pixels x inside regions R generated by the initial watershed segmentation, and $\upsilon_E(\cdot)$ accounts for edge strength between adjacent regions ξ (connected regions).

The original IRGS is applied to the whole SAR scene. However, with the spatial size increases, the incidence angle varies considerably across the SAR scene, causing statistical non-stationarities for each class. Luckily, these non-stationarities only pose issues at larger scales. This characteristic inspires processing SAR images on smaller scales to mitigate incidence angle variation-related challenges. Therefore, a segmentation strategy called 'glocal' [50], which combines local details and global statistics, is combined with original IRGS and applied to SAR scenes to generate superpixels.

First, the entire scene is divided into smaller regions, called autopolygons [50], using a modified watershed algorithm [70] with seeds from a 12×12 grid. Then, IRGS is applied to each autopolygon individually to produce an oversegmentation, which is the local step of the glocal. Lastly, the global step glues the oversegmentation regions across the full scene, creating larger regions. Such a glocal strategy can provide robust segmentation and divide the full scene into homogeneous regions, called superpixels, with high homogeneity and compactness. Consequently, each superpixel is regarded as a homogeneous entity, representing a node in the graph.

Given a dual-pol SAR image $X = \{x_1, x_2, \cdots, x_N\} \subset \mathbb{R}^{2\times N}$, where $x_i = \{x_{i,HH}, x_{i,HV}\}$ denotes the HH and HV polarizations for pixels in terms of intensity, while N is number of pixels in the SAR image. IRGS segmentation algorithm divided SAR image X into a superpixel set $S = \{s_1, s_2, \cdots, s_M\}$, where $s_m = \{x_m^1, x_m^2, \cdots, x_m^{n_m}\}$ $(m = 1, 2, \cdots, M)$ is the m^{th} superpixel. n_m represents the number of pixels in the superpixel s_m, and M is the number of superpixels.

In summary, introducing superpixels reduces the computational cost and the processing time of sea ice classification. It also preserves the local structure between homogeneous regions, as the adjacent superpixels with similar features are likely to have the same ice type. Furthermore, the process of generating superpixels and assigning labels to them mirrors the pipeline of creating ice charts: dividing the SAR scene into polygons and determining the ice concentration and stage of development within them.

4.2.2 CNN-based feature extractor

Upon generating superpixels through the glocal-based IRGS, feature vectors for each superpixel are extracted. Applying the CNN model for feature extraction in the proposed IceGCN is ideal due to its capacity for learning high-level spatial relationships via hierarchical convolutional layers. Generally, a CNN's performance and depth exhibit a positive correlation. However, as the number of layers increases, minor gradient changes may amplify during backpropagation, resulting in exploding and vanishing gradients. Various alternative models, such as the renowned ResNet [122], have been proposed to address performance degradation in classic CNNs.

ResNet introduces the residual block and identity mapping to tackle the performance degradation issues found in traditional CNNs with deeper architectures. To extract adequate high-level features characterizing the sea ice stage of development for each pixel in the SAR image, the depth of the CNN-based feature extractor must be substantial. There are two main residual blocks, basic block and bottleneck. The bottleneck block is chosen to construct the feature extractor backbone because the bottleneck has fewer trainable parameters and demands less computational power [122]. Figure 4.2 illustrates the feature extraction module's architecture and the bottleneck block's structure, as utilized in this

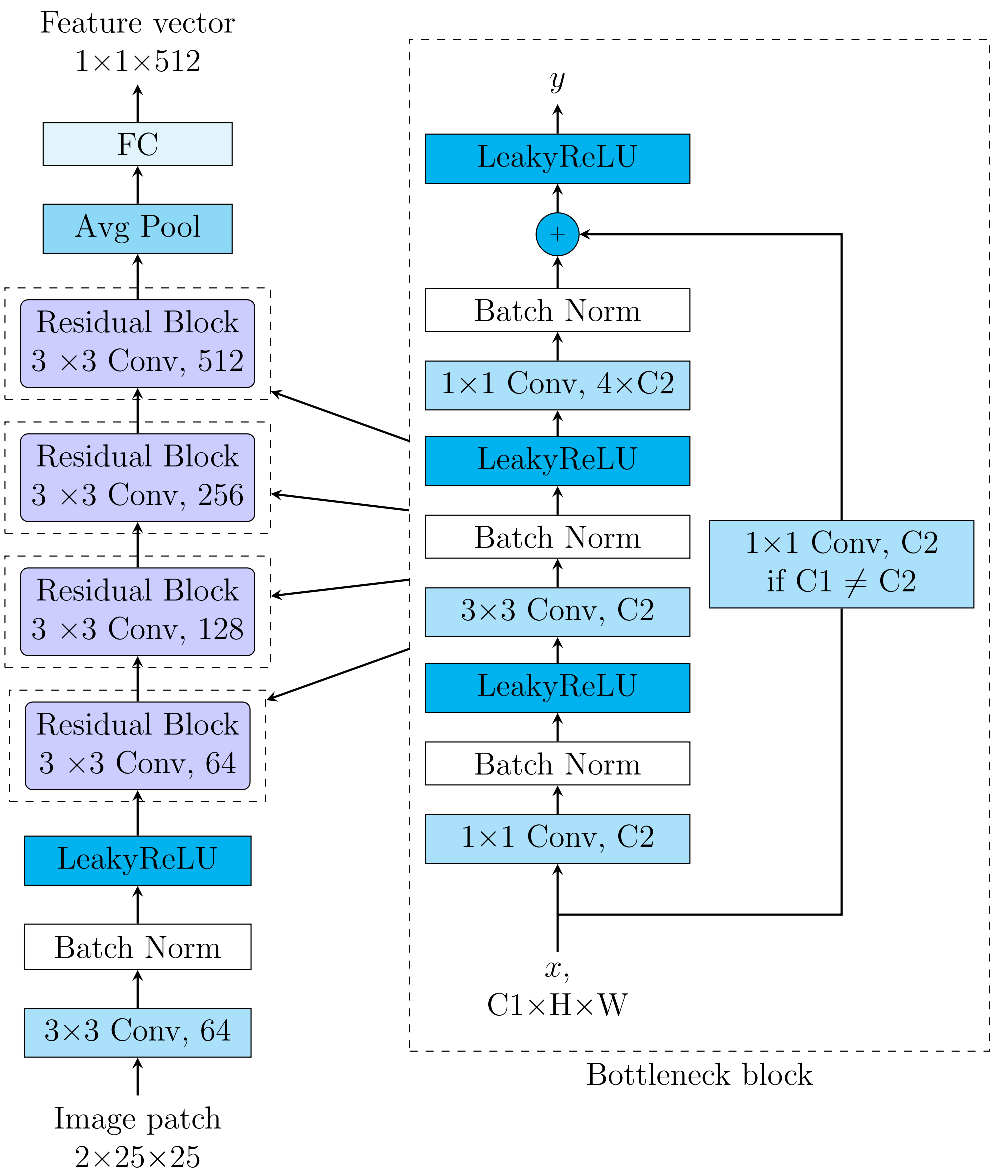


Figure 4.2: Architecture of feature extraction module in IceGCN.

study. This architecture bears similarity to previous research [116], which underwent evaluation for ice-water classification. Owing to the intricate nature of differentiating different sea ice types, the number of output channels of each residual block is quadrupled compared to input channels, enhancing the model's complexity. The loss function for this model is the cross-entropy cost function, which is defined as follows

$$loss = -\frac{1}{n}\sum_{n} y \ln \hat{y} + (1-y)\ln(1-\hat{y}) \tag{4.2}$$

where n is the number of samples in a batch, while y and $\hat{y}$ denote the true and predicted labels, respectively. It's crucial to emphasize that, before integrating with IceGCN, the CNN-based feature extraction module needs to be pre-trained on a distinct dataset.

4.2.3 Graph Construction

Given the generated superpixels and the corresponding feature vectors, the specific steps for constructing the graph are described in this subsection.

Unlike images, which are exhibited in the form of a rectangular lattice in the Euclidean plane, graphs usually have irregular shapes and consist of a set of nodes and connecting edges. Let $G = (V, E)$ represent an undirected graph, where V is a group of vertices (nodes), and E is a set of edges connecting them. A symmetric sparse matrix $\mathbf{A} \in \mathbb{R}^{M\times M}$ called the adjacency matrix (similarity matrix) is used to describe edges between nodes. If any two nodes, v_i and v_j, are connected by an edge $E(v_i, v_j)$ directly, v_i and v_j are considered to be adjacent. $\mathbf{A}_{i,j}$ represents the weight of $E(v_i, v_j)$ between vertices v_i and v_j. The most commonly used definitions of $\mathbf{A}_{i,j}$ are connectivity and distance. However, neither of these can sufficiently address the relationship between the adjacent vertices. Therefore, a similarity function $\mathrm{sim}(i, j)$ is proposed to measure the weight between v_i and v_j. $\mathbf{A}_{i,j}$ in this study is defined as follows,

$$\mathbf{A}_{i,j} = \begin{cases} \mathrm{sim}(i,j)\ , & \text{if } v_i \text{ and } v_j \text{ are adjacent} \\ 0\ , & \text{otherwise} \end{cases} \tag{4.3}$$

where v_i and $v_j \in V$, and M is the number of nodes in the graph. $\mathrm{sim}(i, j)$ is a function that evaluates the similarity between v_i and v_j. Because the Gaussian distribution is often used to model the distribution of backscatter intensity of sea ice in SAR imagery, relative entropy [158] is selected as the function for $sim(i, j)$, which is defined as follows:

$$\begin{aligned} \mathrm{sim}(i,j) = \exp\Bigg[\sum_{\substack{x_i \in v_i \\ x_j \in v_j}} \Bigg(P(x_i) \log \frac{P(x_i)}{P(x_j)} + \\ P(x_j) \log \frac{P(x_j)}{P(x_i)} \Bigg) / 2 \Bigg] \end{aligned} \tag{4.4}$$

where $P(x_i)$ and $P(x_y)$ are the probability distributions of vertices v_i and v_j. Since we assume the backscatter intensity of different sea ice types in the linear scale obey Gaussian distributions in IRGS segmentation, Equation 4.4 can be rewritten as

$$\begin{aligned} \mathrm{sim}(i,j) &= \exp\Bigg[\Bigg(\log(\frac{\sigma_j}{\sigma_i}) + \frac{\sigma_i^2 + (\mu_i - \mu_j)^2}{2\sigma_j^2} - \frac{1}{2} + \\ &\qquad \log(\frac{\sigma_i}{\sigma_j}) + \frac{\sigma_j^2 + (\mu_j - \mu_i)^2}{2\sigma_i^2} - \frac{1}{2} \Bigg) / 2 \Bigg] \\ &= \exp\Bigg(\frac{\sigma_i^2 + (\mu_i - \mu_j)^2}{4\sigma_j^2} + \frac{\sigma_j^2 + (\mu_j - \mu_i)^2}{4\sigma_i^2} - \frac{1}{2} \Bigg) \end{aligned} \tag{4.5}$$

where μ_i, σ_i, μ_j, and σ_j are means and variances of v_i and v_j, respectively.

4.2.4 Graph Convolutional Network

GCNs are inherently suited for processing graph-structured data, which renders them an ideal candidate for application to segmentation results that typically manifest in irregular shapes. Beyond their adaptability to non-Euclidean data, GCNs possess the capacity to learn spatial features and contextual relationships simultaneously. The convolutional operations within GCNs are designed to incorporate the features of adjacent nodes, effectively capturing both spatial patterns and contextual information among superpixels. These capabilities facilitate a comprehensive understanding of the underlying structure within the segmentation results. Therefore, GCN is chosen as the backbone of the proposed method for sea ice classification.

After constructing the graph and calculating the feature vectors of vertices, the unlabeled vertices are classified using the label information propagated from the limited labeled ones. A graph convolutional network [155] is applied for label propagation in this study. The computation inside a basic graph convolutional layer is given by:

$$\mathbf{X}^{l+1} = \mathrm{act}(\mathcal{L}\mathbf{X}^l\mathbf{W}^l) \tag{4.6}$$

where $\mathbf{X}^l$ and $\mathbf{X}^{l+1}$ are the input and output of the l-th layer, respectively. $\mathbf{W}^l$ denotes the learnable weight matrix of l-th layer. $\mathrm{act}(\cdot)$ represents the activation function, which is the rectified linear unit (ReLU) in the proposed model. $\mathcal{L}$ represents the combinatorial Laplacian matrix, which is defined as

$$\mathcal{L} = \mathbf{D} - \mathbf{A} \tag{4.7}$$

Here, $\mathbf{D}$ is the diagonal degree matrix of $\mathbf{A}$ where $\mathbf{D}_{i,i} = \sum_j \mathbf{A}_{i,j}$. The introduction of $\mathbf{D}$ adds features of the node itself to the computation when summing up feature vectors of adjacent nodes. However, the combinatorial Laplacian matrix is usually unnormalized, and therefore the impact of nodes with more neighbors will be amplified when multiplying $\mathcal{L}$ with all the feature vectors of adjacent nodes. Hence, $\hat{\mathbf{A}}$, a variant of $\mathbf{A}$, also known as the symmetric normalized Laplacian matrix, is introduced to replace $\mathcal{L}$

$$\hat{\mathbf{A}} = \tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-\frac{1}{2}} \tag{4.8}$$

where $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}_M$ is the adjacency matrix A with added self-looping. I_M is the identity matrix, and $\tilde{D}$ denotes the degree matrix of $\tilde{A}$. Therefore, (4.6) can be rewritten as:

$$\mathbf{X}^{l+1} = \mathrm{act}(\tilde{\mathbf{A}}\mathbf{X}^l\mathbf{W}^l) \tag{4.9}$$

Let feature set $\mathbf{F} = \{\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_M\} \subset \mathbb{R}^{K\times M}$ represent the extracted features of the superpixel set S, where $K = 512$ is the number of feature extracted. The feature vector $\mathbf{f_m}$ for m^{th} superpixel is defined as the aggregation of the feature vectors of every pixel inside using averaging pooling. The definition of $\mathbf{f}$ is elaborated as follow:

$$\mathbf{f}_m = \frac{1}{n_m}\sum_{i=1}^{n_m} x_m^i \tag{4.10}$$

In summary, the GCN incorporated in the IceGCN consists of two graph convolutional layers and a softmax layer. Assume the output of GCN, a vector of sea ice stage of development for superpixels S, is $\mathbf{y}$, the architecture of GCN can be formulated as follows:

$$\mathbf{y} = \mathrm{softmax}(\hat{\mathbf{A}}\mathrm{ReLu}(\hat{\mathbf{A}}\mathbf{F}\mathbf{W}^0)\mathbf{W}^1). \tag{4.11}$$

4.3 Experiments and Analysis

In this section, the RADARSAT-2 dual-pol dataset used in this study is delineated, along with the training and testing strategies. The performance of the proposed method is then evaluated and compared with benchmark methods.

4.3.1 Data Overview

The dataset used in this study consists of 18 dual-polarized WideScan RADARSAT-2 scenes, which is the same dataset used in a previous study [116]. Table 4.1 shows the scene ID, acquisition date and time, and the orbit of the 18 scenes. Within each scene, the incidence angle ranges from 19°: to 49°. Initially presented in sigma-nought format on a linear scale, the data was subsequently normalized to fit within the [0,1] range, conforming to the input requirements of the CNN model. The size of the original images is around 10000 by 10000 pixels. To enhance computational efficiency, 4×4 block averaging reduces image sizes. Post-downsampling, the nominal pixel spacing is 200 m, which still contains much more details than human-created ice charts.

Upon the generation of superpixels, as delineated in Section 4.2.1, a random subset was chosen to serve as training and testing samples. These samples were subsequently labeled in accordance with the ice chart classifications provided by an analyst from the Canadian Ice Service (CIS). For the purposes of this study, sea ice has been categorized into four distinct types: multi-year ice (MYI), first-year ice (FYI), young ice (YI), and open water (OW).

4.3.2 Sea Ice Appearance in the dataset

The appearance of sea ice in SAR imagery depends on many factors, including environmental conditions, SAR imaging parameters, and the characteristics of the sea ice itself [159]. Figure 4.3 shows patch samples of different ice types in HH and HV polarized scenes of the dataset used in this study. Since the SAR sensor transmits horizontal linear microwave and receives both horizontal and vertical returns, the co-polarization (HH) backscatter typically is higher (brighter in the scene) than that of cross-polarization (HV) [160].

OW presents as relatively dark in both polarizations and even more so in the HV images. This darker appearance is attributed to the sea's reflective nature, which tends to scatter most of the microwave signals away from the SAR antenna. Conversely, YI exhibits

Table 4.1: List of SAR scenes used in this work.

Dataset	SceneID	SAR Acquisition Date (M/D/Y)	Acquisition Time UTC (hh:mm:ss)	Ascending (A)/ Descending (D)
Dataset-1	20100623_041255	06/23/2010	04:12:55	A
	20100629_163326	06/29/2010	16:33:26	D
	20100721_173208	07/21/2010	17:32:08	D
	20100730_162908	07/30/2010	16:29:08	D
	20100807_173610	08/07/2010	17:36:10	D
	20100816_163329	08/16/2010	16:33:29	D
	20100907_035614	09/07/2010	03:56:14	A
	20100909_163321	09/09/2010	16:33:21	D
	20101003_163324	10/03/2010	16:33:24	D
	20101021_041325	10/21/2010	04:13:25	A
	20101114_041304	11/14/2010	04:13:04	A
	20101206_015139	12/06/2010	01:51:39	A
	20101214_025725	12/14/2010	02:57:25	A
Dataset-2	20100418_163315	04/18/2010	16:33:16	D
	20100426_040439	04/26/2010	04:04:39	A
	20100510_035620	05/10/2010	03:56:20	A
	20100524_034756	05/24/2010	03:47:56	A
	20101027_025726	10/27/2010	02:57:26	A

more complex surface structures compared to other types of sea ice. This complexity arises from dynamic processes such as collisions and fracturing, which are driven by the forces of ocean currents and wind. As shown in Figure 4.3(c) and 4.3(d), there can be numerous fissures and ridges presented in a sample of YI. Fissures and ridges are still noticeable for first-year ice, and ice floes start to appear in the scene. Compared with other sea ice types, MYI is recognizable in both polarized scenes, particularly in HV images. As sea ice ages, it naturally drains brine to decrease its salinity. Such a change allows the C-band SAR signal to penetrate the ice and generate volume scattering that characterizes the older ice.

In addition, influences such as temperature and humidity changes may create melt ponds that obscure ice completely. The presence of wet snow can significantly alter the microwave penetration depth, thereby changing the characteristic backscatter of the ice. Wind over open water increases its roughness and backscatter characteristics [161]. Lower temperature contributes to higher backscatter, while the presence of wet snow caused by higher temperature usually reduces it [159].

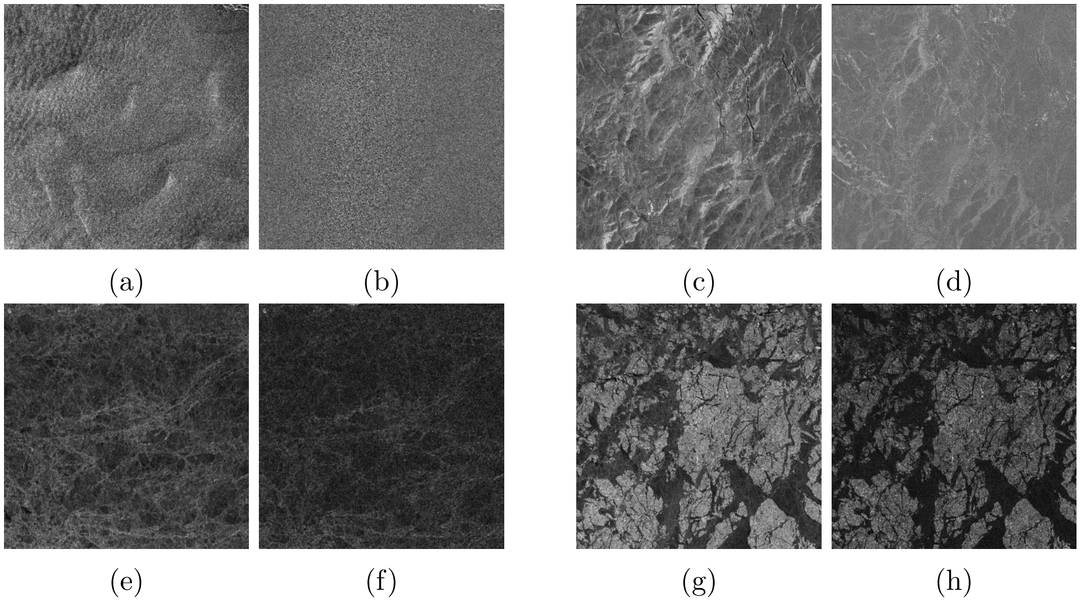


Figure 4.3: Different stages of development of sea ice in patches cropped from HH an HV polarized scenes in the same location. Open water in HH (a) and HV (b). Yong ice in HH (c) and HV (d). First-year ice in HH (e) and HV (f). Multi-year ice in HH (g) and HV (h).

4.3.3 Experimental Setup

Experiments are conducted on a workstation with an Intel(R) Core(TM) i9-9920X CPU @ 3.50GHz $\times$ 24 threads, 128 GB RAM, and three GeForce GTX 2080Ti GPUs with 12 GB of memory. All the deep learning models are implemented using PyTorch, and the IRGS segmentation algorithm is delivered by the MAp-Guided Ice Classification (MAGIC) system [74].

Since the feature extraction module in IceGCN requires a separate dataset for pre-training, we split the 18-scene dataset into two subsets, Dataset-1 and Dataset-2, to avoid using samples from the same scene for training the feature extraction module and evaluating the proposed IceGCN. The feature extraction module is trained, validated, and tested on Dataset-1, which consists of 13 scenes. Dataset-2 is comprised of the remaining five scenes and is utilized for evaluating the IceGCN and comparison methods. The number of superpixels for training and testing of Dataset-2 is provided in Table 4.2.

As delineated in Section 4.3.1, the generation of training and testing samples is conducted at the superpixel level. Following the creation of superpixels for each SAR scene, a scheme is implemented to select them randomly. These selected superpixels are then annotated in alignment with the associated sea ice charts. It is important to recognize that these ice charts serve merely as an approximate guide to the distribution of sea ice within a given scene. Consequently, expert judgment was also employed during the labeling process to enhance the accuracy of the training and testing samples. This approach, while informed, does acknowledge the potential for discrepancies between the labeled superpixels and the information presented in the sea ice charts, particularly in regions proximate to boundaries where the delineation is more prone to ambiguity.

Due to the nature of semi-supervised models, when evaluating a scene, C, IceGCN is trained using samples from Dataset-2, while its feature extraction module undergoes pre-training on Dataset-1. Subsequently, the model is evaluated using testing samples from C. Considering that benchmark methods are fully supervised, they are initially trained on Dataset-1 and then fine-tuned with training samples from Dataset-2, ensuring a just and equitable training strategy for all methods examined in the experiments.

Table 4.2: Number of superpixels for training and testing in Dataset-2.

Scene ID	Ice type	# of train	# of test
20100418	FYI	118	234
	MYI	204	395
20100426	FYI	243	517
	MYI	64	513
20100510	OW	27	49
	FYI	68	151
	MYI	208	452
20100524	OW	34	68
	FYI	110	213
	MYI	135	259
20101027	OW	96	157
	YI	172	293
	MYI	5	11

The performance of random forest (RF), ResNet, IRGS-RF, and IRGS-ResNet in sea ice classification was explored in Chapters 2 and 3. However, their efficacy when trained with a limited number of labeled samples has not yet been assessed. Consequently, these models are designated as baselines for comparative analysis in the current study. Despite the fact that RF and ResNet rely on features extracted from limited receptive fields and may not

Table 4.3: GLCM parameters used to train the RF model in the study.

Receptive field (pixels)	**Step size (pixels)**
5×5	1
11×11	1
25×25	1
25×25	5
51×51	5
51×51	10
51×51	20
101×101	10
101×101	20

capture contextual information on larger scales, their widespread use in the classification of remote sensing images warrants their inclusion in this comparative study.

Several studies have endeavored to amalgamate spatial features with contextual information to differentiate between sea ice and water [141, 142, 48]. Nevertheless, research focusing on the discrimination of various sea ice types employing this combined approach remains scarce. Zhu *et al.* [45] proposed a novel method that harnesses a conditional random field (CRF) applied to the classification outputs obtained from an SVM. This method did result in enhanced classification accuracy in comparison to SVM alone and other similar models. It was particularly effective in restoring details along sea ice boundaries and refining some of the SVM's classification inaccuracies. However, the application of CRF across an entire SAR scene means that the SVM-CRF approach is susceptible to the distortive effects of speckle noise and variations in the incidence angle. These susceptibilities can induce new classification errors, especially at the near and far ranges of SAR scenes. Due to these limitations, the overall efficacy of SVM-CRF is anticipated to be comparable to that of IRGS-RF. Consequently, SVM-CRF has not been included in the set of models selected for the comparative experiments in this research.

Although RF and ResNet offer certain advantages, it is anticipated that IRGS-RF and IRGS-ResNet will outperform them, achieving higher accuracy both qualitatively and quantitatively. However, the contextual features learned by IRGS-RF and IRGS-ResNet are restricted to local vicinities. This limitation could potentially compromise their performance in SAR scenes characterized by complex ice conditions. In contrast, the GCN module within IceGCN possesses the inherent capability to incorporate contextual information on a global scale through feature propagation among superpixels. This attribute theoretically positions IceGCN to deliver superior classification results when compared to the aforementioned models.

The ResNet model shares the same architecture as the feature extraction module in IceGCN, except a softmax layer is added to ResNet for predicting. To train the RF-based model, GLCM features are extracted, which consist of angular second moment, contrast, correlation, dissimilarity, entropy, homogeneity, inverse moment, mean, and standard deviation with the relative angle at 0°, and the receptive field and step sizes listed in Table 4.3. The hyperparameters of RF are set as follows using a cross-validation-based grid search [116]: the number of trees=250, max depth=10, and minimum samples per leaf=2.

4.3.4 Experimental Results

The experimental results are presented in two ways. First, the classification accuracy of RF, IRGS-RF[116], ResNet, IRGS-ResNet[162], and IceGCN is computed based on the correctly classified sample pixel count, providing a quantitative performance evaluation of the proposed methods against benchmark methods. Second, a visual analysis of the classification maps generated by RF, IRGS-RF, ResNet, IRGS-ResNet, and IceGCN offers insights into the qualitative performance of IceGCN and the comparative methods for real-world operational applications.

The quantitative results obtained by RF, IRGS-RF, ResNet, IRGS-ResNet, and IceGCN are reported in terms of classification accuracy in Table 4.4, where the highest accuracies among all methods for each scene are highlighted in bold for each row. In general, the proposed IceGCN achieves the best classification performance with an average/overall accuracy of 95.54%, where the accuracies of each scene are 91.20%, 97.52%, 96.39%, 95.61%, and 95.77%, respectively. Unsurprisingly, the performance of IceGCN is generally superior to that of the comparison methods because both local and global spatial information is utilized for classification. IceGCN outperforms other methods in all scenes in the test dataset (Dataset-2), demonstrating the robustness of the proposed method.

Comparing the two benchmark methods shows that the ResNet outperforms RF in three out of four scenes. In detail, RF acquires an overall accuracy of 80.00%, where the accuracies of each scene are 68.35%, 87.08%, 82.46%, 83.57%, and 72.44%, respectively. Although RF delivers relatively good results in distinguishing FYI and OW, it misclassifies MYI to FYI in most scenes, especially the scene captured on April 18, 2010. In contrast, ResNet achieves an overall accuracy of 84.56% without a significant performance drop in any sea ice type. The classification accuracies of each scene are 83.30%, 82.62%, 87.42%, 85.94%, and 86.12%, respectively. The quantitative results indicate that ResNet is a better method when only considering pixel-wise classifiers for operational sea ice monitoring. For

Table 4.4: Classification accuracy of RF, IRGS-RF, ResNet, IRGS-ResNet, and IceGCN.

Scene ID	Category	RF	IRGS-RF	ResNet	IRGS-ResNet	IceGCN
20100418	FYI	95.32%	**95.83%**	79.37	79.32%	85.24%
	MYI	52.37%	57.62%	84.26%	85.65%	**94.73%**
	Overall	68.35%	71.83%	83.30%	83.04%	**91.20%**
20100426	FYI	80.27%	81.47%	75.81%	76.76%	**98.31%**
	MYI	93.94%	93.35%	89.48%	90.38%	**96.72%**
	Overall	87.08%	87.39%	82.62%	83.54%	**97.52%**
20100510	OW	96.16%	96.30%	94.85%	95.69%	**96.74%**
	FYI	94.68%	95.05%	80.91%	83.70%	**98.66%**
	MYI	76.89%	80.25%	88.79%	87.61%	**95.59%**
	Overall	82.46%	84.88%	87.42%	88.31%	**96.39%**
20100524	OW	96.58%	**98.97%**	92.65%	92.40%	95.31%
	FYI	92.46%	94.48%	80.72%	86.41%	**99.24%**
	MYI	72.84%	76.54%	88.47%	87.96%	**92.77%**
	Overall	83.57%	86.44%	85.94%	87.90%	**95.61%**
20100027	OW	90.75%	94.58%	93.98%	**95.03%**	94.93%
	YI	63.83%	70.64%	81.97%	90.52%	**97.37%**
	MYI	40.46%	34.04%	**84.33%**	75.78%	65.15%
	Overall	72.44%	77.92%	86.12%	91.70%	**95.77%**
Overall accuracy		80.00%	82.47%	84.56%	86.09%	**95.54%**

two-stage models, IRGS-ResNet achieves better classification performance than IRGS-RF, with an overall accuracy of 86.09%.

Since quantitative classification rates are calculated using limited labeled data, visually inspecting classification maps provides an intuitive basis for performance evaluation. Classification results of the scene acquired on April 18, 2010 are presented in Figure 4.4. Because RF and ResNet do not have any prior knowledge about the testing scene, they treat each scene equally. RF and ResNet misclassify some regions into YI, which does not appear in this scene. In contrast, IceGCN does not suffer from this problem owing to the prior knowledge introduced by the limited labeled sample of the scene, i.e., human interactions. Figure 4.4(d) demonstrates that significant misclassifications of MYI are presented in the classification maps produced by RF, especially in the upper part of the scene.

Although IRGS-RF enhances the quality of the sea ice map in Figure 4.4(e) by properly classifying regions near sea ice boundaries and suppressing misclassification of YI, the classification maps produced by RF-based models have the worst visual quality. This

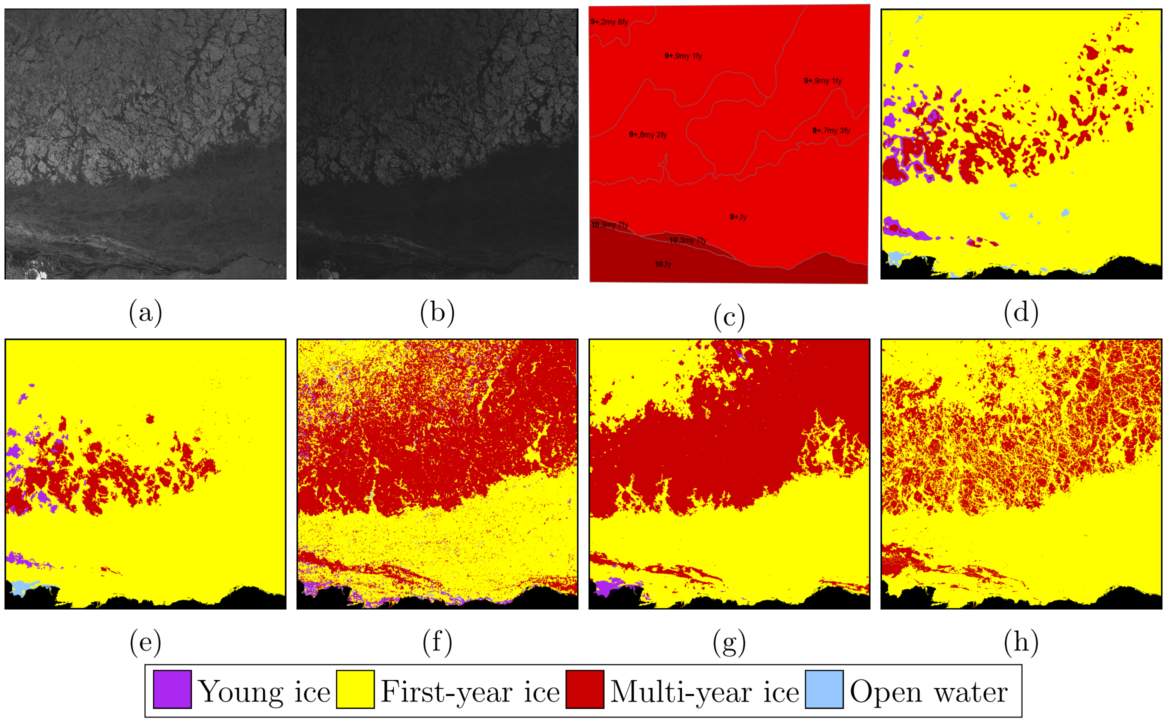


Figure 4.4: Classification results for the scene obtained on April 18, 2010 (scene ID: 20100418). (a) HH polarization (b) HV polarization (c) Ice chart (d) RF (e) IRGS-RF (f) ResNet (g) IRGS-ResNet (h) IceGCN.

reveals that the RF with handcrafted features is inadequate for discriminating sea ice types in complex scenes. The classification map of ResNet in Figure 4.4(f) is improved compared with that of RF. However, large-scale salt-and-pepper-like classification errors appear over the whole image, which is caused by the fact that ResNet does not capture contextual information of the center pixel. After integrating with segmentation results, IRGS-ResNet eliminates most salt-and-pepper classification noise in Figure 4.4(g). However, the FYI formed from frozen leads among MYI is missing since it is not preserved in the results of ResNet. Compared with other methods, the proposed IceGCN removes salt-and-pepper classification noise while preserving the edges and boundaries between different sea ice types that are accurately labeled.

Figure 4.5 shows the results of the scene collected on May 24, 2010. The misclassifications of YI arise again in the classification maps of RF and ResNet. The salt-and-pepper classification noise is severe in the result of ResNet, though the outlines of major sea ice regions are mostly preserved. These similar classification errors are also present in the

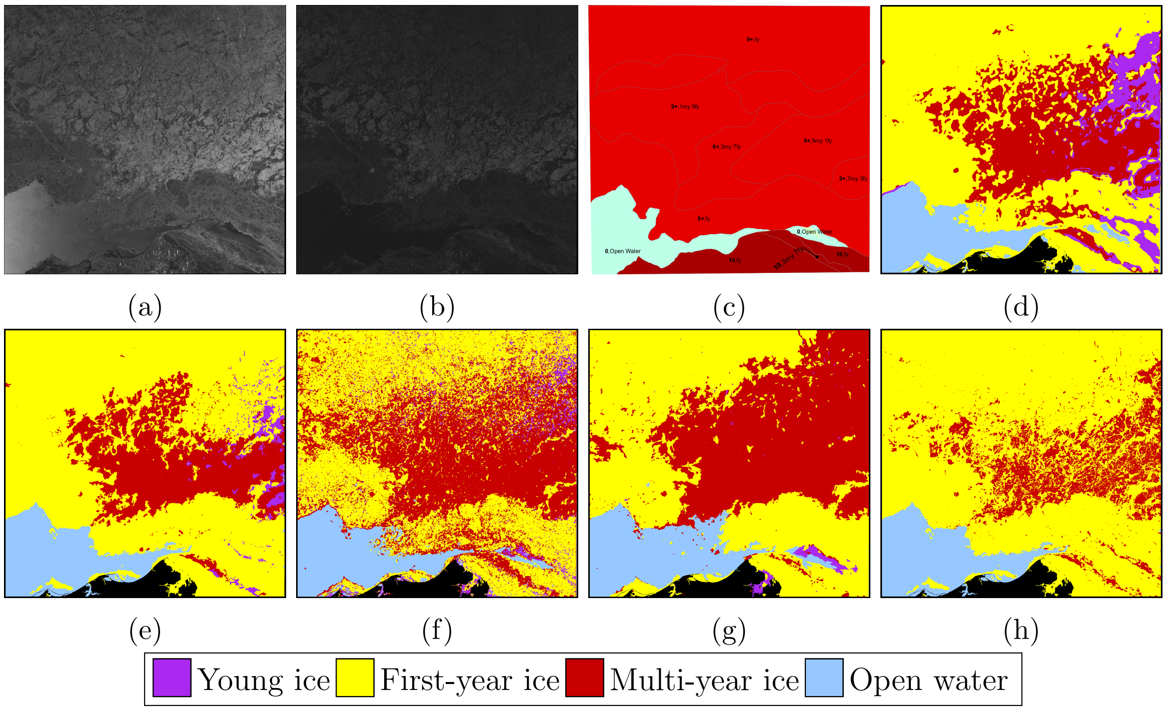


Figure 4.5: Classification results for the scene obtained on May 24, 2010 (scene ID: 20100524). (a) HH polarization (b) HV polarization (c) Ice chart (d) RF (e) IRGS-RF (f) ResNet (g) IRGS-ResNet (h) IceGCN.

classification map of IRGS-ResNet since the proposed method's infrastructure is based on ResNet. Fortunately, the superpixels generated by IRGS suppress the salt-and-pepper classification noise and contribute to a smoother, more appealing result than that of ResNet. The result of RF containing much less noise-like misclassification is attributed to the multi-scale spatial context information captured in the GLCM features.

Figure 4.6 presents another instance underscoring IceGCN's superiority. Captured during the freeze-up season on October 27, 2010, this scene exhibits considerable incidence angle variation. Notably, water, more susceptible to incidence angle variation than sea ice, appears markedly brighter at lower incidence angles (left side of the scene) compared to higher incidence angles (right side of the scene). The ice-covered area features fissures, ridges, and freezing-up leads, complicating the differentiation of YI and FYI for benchmark methods. RF faces challenges in distinguishing YI, misclassifying nearly half of it as FYI, although IRGS-RF slightly mitigates misclassification. ResNet and IRGS-ResNet more effectively classify YI but falter in predicting MYI due to the sparse presence of MYI in

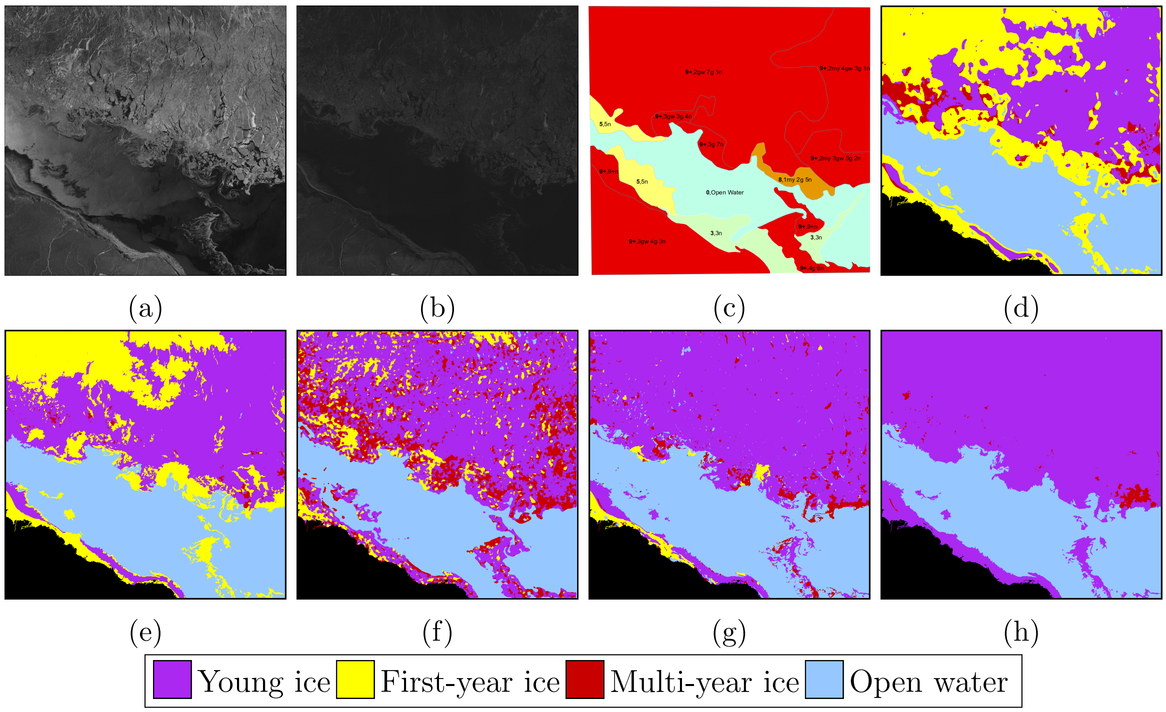
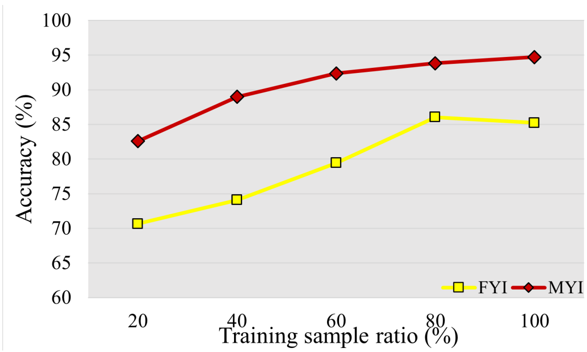


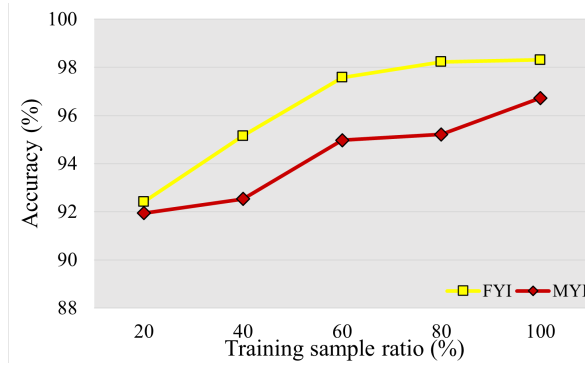
Figure 4.6: Classification results for the scene obtained on October 27, 2010 (scene ID: 20101027). (a) HH polarization (b) HV polarization (c) Ice chart (d) RF (e) IRGS-RF (f) ResNet (g) IRGS-ResNet (h) IceGCN.

the scene and the similar backscatter characteristics of MYI and YI. In contrast, IceGCN attains significantly higher accuracy in classifying YI, owing to human interaction that eliminates FYI from training data. Nonetheless, the limited MYI samples—five for training and eleven for testing—result in a suboptimal classification accuracy of 65.15% for MYI. Despite this, IceGCN still surpasses the benchmark methods by a substantial margin in overall accuracy.

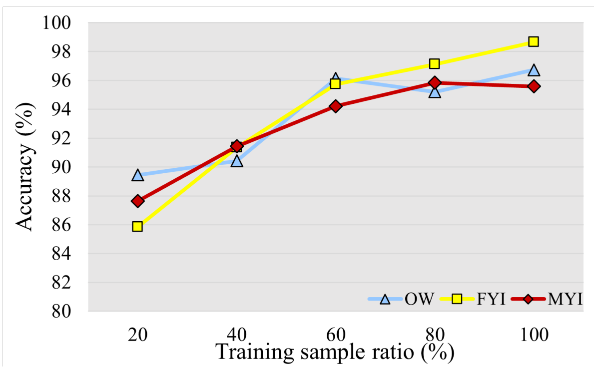
The performance of IceGCN is also assessed using various training sample quantities. Training samples per class are randomly chosen, ranging from 20% to 100% of available samples, in increments of 20%. Figure 4.7 illustrates the classification accuracies. A marked enhancement in classification accuracy is discernible as the training sample ratio advances from 20% to 60%. Beyond the 60% threshold, the increasing trend plateaus—except for MYI in scene 20101027, where an extremely limited number of samples exist. These findings demonstrate that IceGCN can produce remarkable classification outcomes even when constrained by a limited pool of labeled samples.



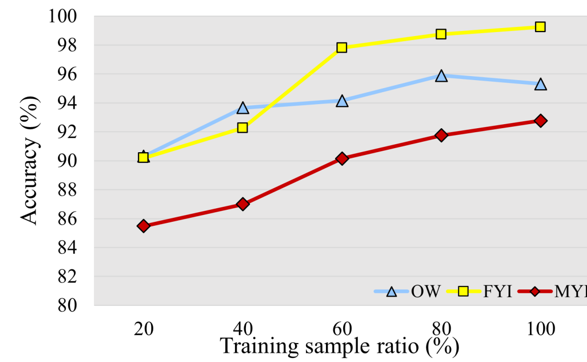
(a). Scene: 20100418



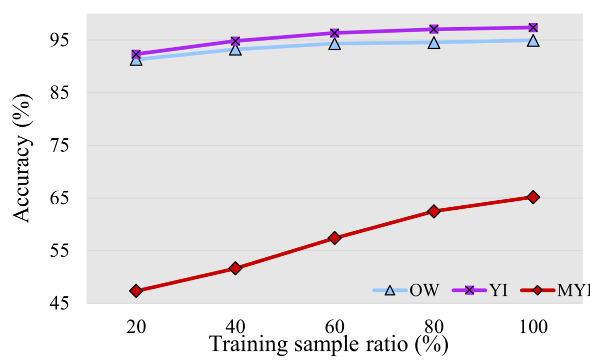
(b). Scene: 20100426



(c). Scene: 20100510



(d). Scene: 20100524



(e). Scene: 20101027

Figure 4.7: Classification accuracy of IceGCN on Dataset-2 with different training sample ratios.

4.4 Summary

The IRGS segmentation-generated superpixels maintain homogeneity, preserving edges and boundaries among various sea ice types. Additionally, GCN propagates feature information from labeled nodes, classifying unlabeled nodes in the graph. Consequently, a graph-based method, IceGCN, is demonstrated to improve the sea ice classification accuracy using RADARSAT-2 duel-pol scenes. IceGCN consists of three primary components: superpixel generation, feature extraction, and graph convolution. The experimental results demonstrate that the proposed method outperforms the other comparison methods in both quantitative and qualitative assessments.

Due to the semi-supervised nature of IceGCN, human involvement is necessary to supply initial training samples for each SAR image. This process constrains the distribution of sea ice within the scene and minimizes the misclassification of non-existent sea ice types. This characteristic is similar to the procedure used by human experts to generate sea ice charts. Experimental findings indicate that classification maps created by IceGCN are more natural, precise, and potentially better suited for operational use than those generated by benchmark methods. Furthermore, IceGCN demands only a limited number of training samples to predict a SAR scene. Typically, the labeling process takes an ice expert under five minutes, resulting in a substantial time reduction compared to the prevailing human-centric pipeline.

In future work, more ablation studies can be conducted to investigate the contributions of the components in IceGCN. Potential investigations may involve testing different superpixel generation and feature extraction algorithms. Moreover, IceGCN currently focuses on local relations for adjacency matrix construction, so exploring the integration of local and global information for graph construction may prove insightful.

Chapter 5

Conclusion and Future work

5.1 Summary of Contributions and Results

The primary objectives of this thesis are to design and develop innovative machine learning frameworks, with a special emphasis on sea ice classification for operational applications, utilizing SAR imagery. In light of identified research gaps from existing studies, the integration of spatial and contextual features emerges as a promising avenue to fulfill these objectives. Hence, this thesis endeavors to unravel methods for the extraction and utilization of spatial and contextual features in SAR imagery to enhance sea ice classification.

Drawing inspiration from the methodologies employed by the Canadian Ice Service in generating ice charts, our approach to modeling sea ice classification parallels their operational pipeline. This methodology incorporates an initial unsupervised segmentation to delineate homogeneous regions, which are subsequently labeled using advanced learning algorithms. Owing to the proven efficacy of the IRGS segmentation algorithm, this thesis predominantly focuses on the development of labeling techniques and their seamless integration with the segmentation process.

Despite the popularity of deep-learning-based models in the remote sensing domain and their validated performance, the substantial demand for extensive labeled samples hampers their application in distinguishing diverse sea ice types at the pixel level. Consequently, a classification method that depends on spatial features extracted by human-engineered metrics, denoted as IRGS-RF, is introduced in Chapter 2. This method employs a random forest classifier to provide initial pixel labels, while IRGS captures contextual information within the SAR scene. An energy minimization framework is then devised to refine the initial results by combining outcomes from both IRGS and RF. The robustness of this strategy is corroborated through its application to datasets exhibiting significant temporal variations, including melting seasons and shifts in incidence angle in dual-polarization SAR data.

In contrast to the intricate task of classifying different sea ice types, obtaining label samples for the differentiation of sea ice from open water is comparatively straightforward. Thus, Chapter 3 proposes a convolutional neural network based approach, termed IceNet, which proficiently discriminates between sea ice and open water. The integration of a regional pooling layer allows for the exploitation of spatial features discerned through labeling, alongside the contextual information obtained from segmentation. Evaluations conducted on a dataset comprising 21 RADARSAT-2 scenes of the Beaufort Sea from 2010 demonstrate the superior accuracy of IceNet over both benchmark and referenced methodologies. Notably, IceNet ensures high precision while preserving the integrity of ice-water boundaries and minimizing misclassifications.

Although the IRGS-RF and IceNet models exhibit notable efficacy and reliability, the contextual features they leverage are heavily contingent on the performance of the IRGS segmentation, which is confined to localized regions as opposed to the entire scene. Furthermore, the scarcity of datasets with adequate labeled samples continues to preclude the use of fully-supervised deep learning models for sea ice classification. Addressing these constraints and the integration challenge, Chapter 4 introduces a semi-supervised graph convolutional network based model, IceGCN. This novel method synergizes local spatial features captured via CNN with global spatial and contextual features learned through GCN, achieving enhanced sea ice classification performance. Significantly, IceGCN has proven to enhance classification accuracy, both quantitatively and qualitatively, even with a limited quantity of training samples.

In summation, the salient contributions of this thesis are as follows:

- Substantial improvement in sea ice classification accuracy from both qualitative and quantitative standpoints.
- Creation of sea ice maps with pixel-level precision and clearly demarcated boundaries between water and various sea ice types.
- First attempt to inherently integrate unsupervised segmentation and deep-learning labeling for distinguishing sea ice in full-scene SAR imagery.
- Implementation of a GCN-based model that facilitates human-interactive operational sea ice monitoring.

- An exhaustive evaluation of the numerical accuracy, quality of the generated maps, and computational efficiency of the proposed methodologies through a leave-one-out cross-validation strategy.

5.2 Future Work

Although we have validated the performance of our proposed methods across a range of seasons in the Beaufort Sea, comprehensive validation demands more extensive datasets collected from various locations and across different years. Future studies can extend to testing the methods with data acquired from new locations and assessing their generalizability. Additionally, we can explore the potential of using circularly polarized (CP) SAR data from RCM, as this data type has shown promising potential for improving sea ice classification compared to dual-polarized data [105].

Additionally, comprehensive ablation studies could be performed to assess the individual contributions of various components within the IceGCN model. These investigations might include testing different superpixel generation algorithms, feature extraction techniques, and exploring the fusion of local and global information for constructing more efficient adjacency matrices.

References

[1] Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology, "Ice chart colour code standard, version 1.0, 2014," World Meteorological Organization & Intergovernmental Oceanographic Commission, Geneva, Switzerland, Tech. Rep. (WMO TD: 1215b), 2014.

[2] (2022) Ice chart of Davis Strait on 03 july 2022. [Online]. Available: https://ice-glaces.ec.gc.ca/prods/WIS33C/20220703180000_WIS33C_0012165712.pdf

[3] Q. Yu, "Automated SAR sea ice interpretation," Ph.D. dissertation, University of Waterloo, Waterloo, Ont, 2006.

[4] B. Scheuchl, D. Flett, R. Caves, and I. Cumming, "Potential of RADARSAT-2 data for operational sea ice monitoring," *Can. J. Remote Sens.*, vol. 30, no. 3, pp. 448–461, 2004.

[5] A. A. Thompson*, "Overview of the RADARSAT constellation mission," *Can. J. Remote Sens.*, vol. 41, no. 5, pp. 401–407, 2015.

[6] (2016) Manual of ice (MANICE), chapter 5: Ice analysis charts. [Online]. Available: https://www.canada.ca/en/environment-climate-change/services/weather-manuals-documentation/manice-manual-of-ice/chapter-5.html

[7] N. Zakhvatkina, A. Korosov, S. Muckenhuber, S. Sandven, and M. Babiker, "Operational algorithm for ice–water classification on dual-polarized RADARSAT-2 images," *The Cryosphere*, vol. 11, no. 1, pp. 33–46, 2017.

[8] L. P. Bobylev and M. W. Miles, "Sea ice in the arctic paleoenvironments," in *Sea Ice in the Arctic*. Springer, 2020, pp. 9–56.

[9] (2016) Interpreting ice charts. [Online]. Available: https://www.canada.ca/en/environment-climate-change/services/ice-forecasts-observations/publications/interpreting-charts.html

[10] S. Khaleghian, H. Ullah, T. Kræmer, N. Hughes, T. Eltoft, and A. Marinoni, "Sea ice classification of SAR imagery based on convolution neural networks," *Remote Sens.*, vol. 13, no. 9, p. 1734, 2021.

[11] S. Khaleghian, H. Ullah, T. Krmer, T. Eltoft, and A. Marinoni, "Deep semi-supervised teacher-student model based on label propagation for sea ice classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021.

[12] L. Zhang, H. Liu, X. Gu, H. Guo, J. Chen, and G. Liu, "Sea ice classification using TerraSAR-X ScanSAR data with removal of scalloping and interscan banding," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 2, pp. 589–598, 2019.

[13] B. Scheuchl, R. Caves, D. Flett, R. De Abreu, M. Arkett, and I. Cumming, "The potential of cross-polarization information for operational sea ice monitoring," in *Envisat & ERS Symposium*, vol. 572, 2005, pp. 1–7.

[14] K. D. Ward, S. Watts, and R. J. Tough, *Sea clutter: scattering, the K distribution and radar performance*. Stevenage: Institution of Engineering and Technology, 2006.

[15] M. Mäkynen and M. Hallikainen, "Investigation of C-and X-band backscattering signatures of Baltic sea ice," *International Journal of Remote Sensing*, vol. 25, no. 11, pp. 2061–2086, 2004.

[16] W. Dierking, "Mapping of different sea ice regimes using images from Sentinel-1 and ALOS synthetic aperture radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 3, pp. 1045–1058, 2010.

[17] J. P. Gill and J. J. Yackel, "Evaluation of C-band SAR polarimetric parameters for discrimination of first-year sea ice types," *Can. J. Remote Sens.*, vol. 38, no. 3, pp. 306–323, 2012.

[18] S. Singha, M. Johansson, N. Hughes, S. M. Hvidegaard, and H. Skourup, "Arctic sea ice characterization using spaceborne fully polarimetric L-, C-, and X-band SAR with validation by airborne measurements," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 7, pp. 3715–3734, 2018.

[19] L. He, X. He, F. Hui, Y. Ye, T. Zhang, and X. Cheng, "Investigation of polarimetric decomposition for arctic summer sea ice classification using gaofen-3 fully polarimetric sar data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 3904–3915, 2022.

[20] M. Dabboor and M. Shokr, "Assessment of compact polarimetric sar parameters for lake and fast sea ice characterisization," in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2019, pp. 5840–5842.

[21] J. Karvonen, "Virtual radar ice buoys–a method for measuring fine-scale sea ice drift," *The Cryosphere*, vol. 10, no. 1, pp. 29–42, 2016.

[22] J. Lehtiranta, S. Siiriä, and J. Karvonen, "Comparing c-and l-band sar images for sea ice motion estimation," *The Cryosphere*, vol. 9, no. 1, pp. 357–366, 2015.

[23] D. Murashkin, G. Spreen, M. Huntemann, and W. Dierking, "Method for detection of leads from Sentinel-1 SAR images," *Annals of Glaciology*, vol. 59, no. 76pt2, p. 124–136, 2018.

[24] D. A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization," *Can. J. Remote Sens.*, vol. 28, no. 1, pp. 45–62, 2002.

[25] H. Liu, H. Guo, X.-M. Li, and L. Zhang, "An approach to discrimination of sea ice from open water using SAR data," in *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*. IEEE, 2016, pp. 4865–4867.

[26] H. Su, Y. Wang, J. Xiao, and X.-H. Yan, "Classification of MODIS images combining surface temperature and texture features using the support vector machine method for estimation of the extent of sea ice in the frozen Bohai Bay, China," *International Journal of Remote Sensing*, vol. 36, no. 10, pp. 2734–2750, 2015.

[27] W. Tan, "Sea ice mapping in Labrador coast with Sentinel-1 synthetic aperture radar imagery," Master's thesis, University of Waterloo, 2017.

[28] M. S. Mahmud, V. Nandan, S. E. Howell, T. Geldsetzer, and J. Yackel, "Seasonal evolution of L-band SAR backscatter over landfast Arctic sea ice," *Remote Sens. Environ.*, vol. 251, p. 112049, 2020.

[29] S. Singha, A. M. Johansson, and A. P. Doulgeris, "Robustness of SAR sea ice type classification across incidence angles and seasons at L-Band," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–12, 2020.

[30] M. Dabboor, B. Montpetit, S. Howell, and C. Haas, "Improving sea ice characterization in dry ice winter conditions using polarimetric parameters from C-and L-band SAR data," *Remote Sens.*, vol. 9, no. 12, p. 1270, 2017.

[31] M. S. Mahmud, T. Geldsetzer, S. E. L. Howell, J. J. Yackel, V. Nandan, and R. K. Scharien, "Incidence angle dependence of HH-polarized C- and L-band wintertime backscatter over Arctic sea ice," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 11, pp. 6686–6698, 2018.

[32] M.-A. Moen, A. P. Doulgeris, S. N. Anfinsen, A. H. Renner, N. Hughes, S. Gerland, and T. Eltoft, "Comparison of feature based segmentation of full polarimetric SAR satellite sea ice images with manually drawn ice charts," *The Cryosphere*, vol. 7, no. 6, pp. 1693–1705, 2013.

[33] A. Y. Ng and M. I. Jordan, "On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes," in *Advances in neural information processing systems*, 2002, pp. 841–848.

[34] H. Liu, H. Guo, and L. Zhang, "SVM-based sea ice classification using textural features and concentration from RADARSAT-2 dual-pol ScanSAR data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 4, pp. 1601–1613, 2014.

[35] J. Lohse, A. P. Doulgeris, and W. Dierking, "An optimal decision-tree design strategy and its application to sea ice classification from SAR imagery," *Remote Sens.*, vol. 11, no. 13, p. 1574, 2019.

[36] H. Han, J. Im, M. Kim, S. Sim, J. Kim, D.-j. Kim, and S.-H. Kang, "Retrieval of melt ponds on Arctic multiyear sea ice in summer from TerraSAR-X dual-polarization data using machine learning approaches: A case study in the chukchi sea with mid-incidence angle data," *Remote Sens.*, vol. 8, no. 1, p. 57, 2016.

[37] 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, no. 1, pp. 343–364, 2018.

[38] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 645–657, 2017.

[39] R. Ressel, A. Frost, and S. Lehner, "A neural network-based classification for sea ice types on X-band SAR images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 7, pp. 3672–3680, 2015.

[40] W. Song, M. Li, W. Gao, D. Huang, Z. Ma, A. Liotta, and C. Perra, "Automatic sea-ice classification of SAR images based on spatial and temporal features learning," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

[41] H. Boulze, A. Korosov, and J. Brajard, "Classification of sea ice types in Sentinel-1 SAR data using convolutional neural networks," *Remote Sens.*, vol. 12, no. 13, p. 2165, 2020.

[42] W. Song, M. Li, Q. He, D. Huang, C. Perra, and A. Liotta, "A residual convolution neural network for sea ice classification with Sentinel-1 SAR imagery," in *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2018, pp. 795–802.

[43] Y. Ren, H. Xu, B. Liu, and X. Li, "Sea ice and open water classification of SAR images using a deep learning model," in *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2020, pp. 3051–3054.

[44] R. Jobanputra and D. A. Clausi, "Preserving boundaries for image texture segmentation using grey level co-occurring probabilities," *Pattern Recognit.*, vol. 39, no. 2, pp. 234–245, 2006.

[45] T. Zhu, F. Li, G. Heygster, and S. Zhang, "Antarctic sea-ice classification based on conditional random fields from RADARSAT-2 dual-polarization satellite images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 6, pp. 2451–2467, 2016.

[46] R. Ressel, S. Singha, S. Lehner, A. Rösel, and G. Spreen, "Investigation into different polarimetric features for sea ice classification using X-band synthetic aperture radar," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 7, pp. 3131–3143, 2016.

[47] M. Dabboor, B. Montpetit, and S. Howell, "Assessment of the high resolution sar mode of the radarsat constellation mission for first year ice and multiyear ice characterization," *Remote Sens.*, vol. 10, no. 4, p. 594, 2018.

[48] Y. Zhang, T. Zhu, G. Spreen, C. Melsheimer, M. Huntemann, N. Hughes, S. Zhang, and F. Li, "Sea ice and water classification on dual-polarized Sentinel-1 imagery during melting season," *The Cryosphere Discussions*, vol. 2021, pp. 1–26, 2021.

[49] S. Ochilov and D. A. Clausi, "Operational SAR sea-ice image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, p. 4397, 2012.

[50] S. Leigh, Z. Wang, and D. A. Clausi, "Automated ice–water classification using dual polarization SAR satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5529–5539, 2014.

[51] R. G. Onstott and F. Carsey, "SAR and scatterometer signatures of sea ice," *Microwave remote sensing of sea ice*, vol. 68, pp. 73–104, 1992.

[52] J. Lohse, A. P. Doulgeris, and W. Dierking, "Mapping sea-ice types from Sentinel-1 considering the surface-type dependent effect of incidence angle," *Annals of Glaciology*, vol. 61, no. 83, pp. 260–270, 2020.

[53] M. Mäkynen and J. Karvonen, "Incidence angle dependence of first-year sea ice backscattering coefficient in Sentinel-1 SAR imagery over the Kara sea," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 11, pp. 6170–6181, 2017.

[54] A. S. Komarov and M. Buehner, "Detection of first-year and multi-year sea ice from dual-polarization SAR images under cold conditions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 9109–9123, 2019.

[55] W. Song, W. Gao, Q. He, A. Liotta, and W. Guo, "Si-stsar-7: A large SAR images dataset with spatial and temporal information for classification of winter sea ice in Hudson Bay," *Remote Sensing*, vol. 14, no. 1, p. 168, 2021.

[56] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," *Advances in neural information processing systems*, vol. 28, 2015.

[57] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *Journal of big Data*, vol. 8, pp. 1–74, 2021.

[58] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and their applications*, vol. 13, no. 4, pp. 18–28, 1998.

[59] Y.-R. Wang and X.-M. Li, "Arctic sea ice cover data from spaceborne synthetic aperture radar by deep learning," *Earth System Science Data*, vol. 13, no. 6, pp. 2723–2742, 2021. [Online]. Available: https://essd.copernicus.org/articles/13/2723/2021/

[60] R. Saldo, M. B. Kreiner, J. Buus-Hinkler, L. T. Pedersen, D. Malmgren-Hansen, A. A. Nielsen, and H. Skriver, "AI4Arctic / ASIP Sea Ice Dataset - version 2," Tech. Rep., 6 2021. [Online]. Available: https://data.dtu.dk/articles/dataset/AI4Arctic_ASIP_Sea_Ice_Dataset_-_version_2/13011134

[61] J. Buus-Hinkler, T. Wulf, A. R. Stokholm, A. Korosov, R. Saldo, L. T. Pedersen, D. Arthurs, R. Solberg, N. Longépé, and M. Brandt Kreiner, "AI4Arctic sea ice challenge dataset," Nov 2022. [Online]. Available: https://data.dtu.dk/collections/AI4Arctic_Sea_Ice_Challenge_Dataset/6244065/2

[62] Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology, "Sea ice nomenclature," World Meteorological Organization & Intergovernmental Oceanographic Commission, Geneva, Switzerland, Tech. Rep. (WMO No. 259, volume 1 – Terminology and Codes, Volume II – Illustrated Glossary and III – International System of Sea-Ice Symbols, 2014.

[63] A. Stokholm, T. Wulf, A. Kucik, R. Saldo, J. Buus-Hinkler, and S. M. Hvidegaard, "AI4SeaIce: Toward solving ambiguous SAR textures in convolutional neural networks for automatic sea ice concentration charting," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.

[64] L. Zhao, T. Xie, W. Perrie, and J. Yang, "Deep learning-based sea ice classification with Sentinel-1 and AMSR-2 data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023.

[65] Q. Yu and D. A. Clausi, "SAR sea-ice image analysis based on iterative region growing using semantics," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 12, pp. 3919–3931, 2007.

[66] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[67] P. Soille *et al.*, *Morphological image analysis: principles and applications.* Springer, 1999, vol. 2, no. 3.

[68] Y. Y. Boykov and M.-P. Jolly, “Interactive graph cuts for optimal boundary & region segmentation of objects in nd images,” in *Proceedings eighth IEEE international conference on computer vision. ICCV 2001*, vol. 1. IEEE, 2001, pp. 105–112.

[69] S. Geman and D. Geman, “Stochastic relaxation, gibbs distributions, and the bayesian restoration of images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, pp. 721–741, 1984.

[70] Q. Yu and D. A. Clausi, “IRGS: Image segmentation using edge penalties and region growing,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 12, pp. 2126–2139, 2008.

[71] P. Yu, A. Qin, and D. A. Clausi, “Unsupervised polarimetric SAR image segmentation and classification using region growing with edge penalty,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 4, pp. 1302–1317, 2012.

[72] F. Li, D. A. Clausi, L. Xu, and A. Wong, “ST-IRGS: A region-based self-training algorithm applied to hyperspectral image classification and segmentation,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 1, pp. 3–16, 2017.

[73] M. Ghanbari, D. A. Clausi, and L. Xu, “Cp-irgs: A region-based segmentation of multilook complex compact polarimetric sar data,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 6559–6571, 2021.

[74] D. Clausi, A. Qin, M. Chowdhury, P. Yu, and P. Maillard, “MAGIC: MAp-guided ice classification system,” *Can. J. Remote Sens.*, vol. 36, no. sup1, pp. S13–S25, 2010.

[75] L. Vincent and P. Soille, “Watersheds in digital spaces: an efficient algorithm based on immersion simulations,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, pp. 583–598, 1991.

[76] J. Wang, C. Duguay, D. Clausi, V. Pinard, and S. Howell, “Semi-automated classification of lake ice cover using dual polarization RADARSAT-2 imagery,” *Remote Sens.*, vol. 10, no. 11, p. 1727, 2018.

[77] D. BARBER and E. LEDREW, “SAR sea ice discrimination using texture statistics-A multivariate approach,” *Photogrammetric Engineering and Remote Sensing*, vol. 57, no. 4, pp. 385–395, 1991.

[78] J. H. Friedman, “On bias, variance, 0/1—loss, and the curse-of-dimensionality,” *Data mining and knowledge discovery*, vol. 1, no. 1, pp. 55–77, 1997.

[79] M. Verleysen and D. François, "The curse of dimensionality in data mining and time series prediction," in *International work-conference on artificial neural networks*. Springer, 2005, pp. 758–770.

[80] P. M. Granitto, C. Furlanello, F. Biasioli, and F. Gasperi, "Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products," *Chemometrics and Intelligent Laboratory Systems*, vol. 83, no. 2, pp. 83–90, 2006.

[81] L. Xu, J. Li, and A. Brenning, "A comparative study of different classification techniques for marine oil spill identification using RADARSAT-1 imagery," *Remote Sens. Environ.*, vol. 141, pp. 14–23, 2014.

[82] G. S. Chavan, S. Manjare, P. Hegde, and A. Sankhe, "A survey of various machine learning techniques for text classification," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 15, no. 6, pp. 288–292, 2014.

[83] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 832–844, 1998.

[84] R. Shrivastava, H. Mahalingam, and N. Dutta, "Application and evaluation of random forest classifier technique for fault detection in bioreactor operation," *Chemical Engineering Communications*, vol. 204, no. 5, pp. 591–598, 2017.

[85] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[86] H. Wan, X. Luo, Z. Wu, X. Qin, X. Chen, B. Li, J. Shang, and D. Zhao, "Multi-featured sea ice classification with sar image based on convolutional neural network," *Remote Sensing*, vol. 15, no. 16, p. 4014, 2023.

[87] L. Wang, K. A. Scott, and D. A. Clausi, "Sea ice concentration estimation during freeze-up from sar imagery using a convolutional neural network," *Remote Sens.*, vol. 9, no. 5, p. 408, 2017.

[88] J. Lohse, A. P. Doulgeris, and W. Dierking, "Incident angle dependence of Sentinel-1 texture features for sea ice classification," *Remote Sens.*, vol. 13, no. 4, p. 552, 2021.

[89] R. Kwok, "Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018)," *Environmental Research Letters*, vol. 13, no. 10, p. 105005, 2018.

[90] J. C. Stroeve, M. C. Serreze, M. M. Holland, J. E. Kay, J. Malanik, and A. P. Barrett, "The arctic's rapidly shrinking sea ice cover: a research synthesis," *Climatic change*, vol. 110, no. 3, pp. 1005–1027, 2012.

[91] V. C. Khon, I. Mokhov, M. Latif, V. A. Semenov, and W. Park, "Perspectives of northern sea route and northwest passage in the twenty-first century," *Climatic change*, vol. 100, no. 3, pp. 757–768, 2010.

[92] N. Zakhvatkina, V. Smirnov, and I. Bychkova, "Satellite sar data-based sea ice classification: An overview," *Geosciences*, vol. 9, no. 4, p. 152, 2019.

[93] L. White, K. Millard, S. Banks, M. Richardson, J. Pasher, and J. Duffe, "Moving to the RADARSAT constellation mission: Comparing synthesized compact polarimetry and dual polarimetry data with fully polarimetric RADARSAT-2 data for image classification of peatlands," *Remote Sens.*, vol. 9, no. 6, p. 573, 2017.

[94] D. De Lisle, S. Iris, E. Arsenault, J. Smyth, and G. Kroupnik, "RADARSAT Constellation Mission status update," in *EUSAR 2018; 12th European Conference on Synthetic Aperture Radar*. VDE, 2018, pp. 1–5.

[95] P. R. Holland and R. Kwok, "Wind-driven trends in antarctic sea-ice drift," *Nature Geoscience*, vol. 5, no. 12, pp. 872–875, 2012.

[96] H. S. Anderson and D. G. Long, "Sea ice mapping method for seawinds," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 647–657, 2005.

[97] B. Scheuchl, R. Caves, I. Cumming, and G. Staples, "Automated sea ice classification using spaceborne polarimetric sar data," in *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No. 01CH37217)*, vol. 7. IEEE, 2001, pp. 3117–3119.

[98] M. Makynen, A. T. Manninen, M. Simila, J. A. Karvonen, and M. T. Hallikainen, "Incidence angle dependence of the statistical properties of C-band HH-polarization backscattering signatures of the Baltic Sea ice," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 12, pp. 2593–2605, 2002.

[99] W. Lang, P. Zhang, J. Wu, Y. Shen, and X. Yang, "Incidence angle correction of SAR sea ice data based on locally linear mapping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3188–3199, 2016.

[100] F. Gao, X. Wang, Y. Gao, J. Dong, and S. Wang, "Sea ice change detection in SAR images based on convolutional-wavelet neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 8, pp. 1240–1244, 2019.

[101] W. Wenbo, W. Yusong, D. Xue, J. Xiaotong, K. Yida, and W. Xiangli, "Sea ice classification of SAR image based on wavelet transform and gray level co-occurrence matrix," in *2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*. IEEE, 2015, pp. 104–107.

[102] I. De Gelis, A. Colin, and N. Longépé, "Prediction of categorized sea ice concentration from Sentinel-1 SAR images based on a fully convolutional network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021.

[103] M.-A. Moen, S. N. Anfinsen, A. P. Doulgeris, A. Renner, and S. Gerland, "Assessing polarimetric SAR sea-ice classifications using consecutive day images," *Annals of Glaciology*, vol. 56, no. 69, pp. 285–294, 2015.

[104] R. Ressel and S. Singha, "Comparing near coincident space borne C and X band fully polarimetric SAR data for Arctic sea ice classification," *Remote Sens.*, vol. 8, no. 3, p. 198, 2016.

[105] M. Ghanbari, D. A. Clausi, L. Xu, and M. Jiang, "Contextual classification of sea-ice types using compact polarimetric SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 7476–7491, 2019.

[106] L.-K. Soh and C. Tsatsoulis, "Texture analysis of sar sea ice imagery using gray level co-occurrence matrices," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 2, pp. 780–795, 1999.

[107] B. Wang, L. Xia, D. Song, Z. Li, and N. Wang, "A two-round weight voting strategy-based ensemble learning method for sea ice classification of Sentinel-1 imagery," *Remote Sens.*, vol. 13, no. 19, p. 3945, 2021.

[108] X.-M. Li, Y. Sun, and Q. Zhang, "Extraction of sea ice cover by Sentinel-1 SAR based on support vector machine with unsupervised generation of training data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 3040–3053, 2021.

[109] H. Lyu, W. Huang, and M. Mahdianpari, "Sea ice detection from the radarsat constellation mission experiment data," in *2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE, 2021, pp. 1–4.

[110] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 645–657, 2016.

[111] Y. Han, Y. Liu, Z. Hong, Y. Zhang, S. Yang, and J. Wang, "Sea ice image classification based on heterogeneous data fusion and deep learning," *Remote Sens.*, vol. 13, no. 4, p. 592, 2021.

[112] T. Zhang, Y. Yang, M. Shokr, C. Mi, X.-M. Li, X. Cheng, and F. Hui, "Deep learning based sea ice classification with Gaofen-3 fully polarimetric SAR data," *Remote Sens.*, vol. 13, no. 8, p. 1452, 2021.

[113] N. Asadi, K. A. Scott, A. S. Komarov, M. Buehner, and D. A. Clausi, "Evaluation of a neural network with uncertainty for detection of ice and water in SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 1, pp. 247–259, 2020.

[114] Y. Ren, X. Li, X. Yang, and H. Xu, "Development of a dual-attention U-Net model for sea ice and open water classification on SAR images," *IEEE Geoscience and Remote Sensing Letters*, 2021.

[115] J. Chi, J. Bae, and Y.-J. Kwon, "Two-stream convolutional long-and short-term memory model using perceptual loss for sequence-to-sequence Arctic sea ice prediction," *Remote Sens.*, vol. 13, no. 17, p. 3413, 2021.

[116] M. Jiang, D. A. Clausi, and L. Xu, "Sea ice mapping of RADARSAT-2 imagery by integrating spatial contexture with textural features," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022.

[117] A. Luscombe, "Radarsat-2 sar image quality and calibration operations," *Can. J. Remote Sens.*, vol. 30, no. 3, pp. 345–354, 2004.

[118] H. Choi and J. Jeong, "Speckle noise reduction technique for sar images using statistical characteristics of speckle noise and discrete wavelet transform," *Remote Sens.*, vol. 11, no. 10, p. 1184, 2019.

[119] Y.-R. Wang and X.-M. Li, "Arctic sea ice cover data from spaceborne SAR by deep learning," *Earth Syst. Sci. Data Discuss*, pp. 1–30, 2020.

[120] T.-T. Wong, "Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation," *Pattern Recognit.*, vol. 48, no. 9, pp. 2839–2846, 2015.

[121] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012. [Online]. Available: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

[122] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, June 2016, pp. 770–778.

[123] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*. PMLR, 2015, pp. 448–456.

[124] A. F. Agarap, "Deep learning using rectified linear units (relu)," *arXiv preprint arXiv:1803.08375*, 2018.

[125] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," *arXiv preprint arXiv:1505.00853*, 2015.

[126] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026–1034.

[127] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[128] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[129] M. Hoekstra, M. Jiang, D. A. Clausi, and C. Duguay, "Lake ice-water classification of radarsat-2 images by integrating irgs segmentation with pixel-based random forest labeling," *Remote Sens.*, vol. 12, no. 9, p. 1425, 2020.

[130] Z. Ma, Z. Liu, J. Pu, L. Xu, K. Li, L. Wangqu, R. Wu, Y. Ma, Y. Chen, and C. Duguay, "Deep convolutional neural network with random field model for lake ice mapping from sentinel-1 imagery," *International Journal of Remote Sensing*, vol. 42, no. 24, pp. 9351–9375, 2021.

[131] T. Vihma, "Effects of Arctic sea ice decline on weather and climate: A review," *Surveys in Geophysics*, vol. 35, no. 5, pp. 1175–1214, 2014.

[132] D. Perovich, W. Meier, M. Tschudi, S. Hendricks, A. Petty, D. Divine, S. Farrell, S. Gerland, C. Haas, L. Kaleschke *et al.*, "Arctic report card 2020: sea ice," National Oceanic and Atmospheric Administration, Tech. Rep., 2020.

[133] T. A. Moon, M. L. Druckenmiller, R. L. Thoman *et al.*, "2021: Arctic report card," National Oceanic and Atmospheric Administration, Tech. Rep., 2021.

[134] X.-M. Li, Y. Qiu, Y. Wang, B. Huang, H. Lu, M. Chu, H. Fu, and F. Hui, "Light from space illuminating the polar silk road," *Int. J. Digit. Earth*, vol. 15, no. 1, pp. 2028–2045, 2022.

[135] N. P. Walker, K. C. Partington, M. L. Van Woert, and T. L. Street, "Arctic sea ice type and concentration mapping using passive and active microwave sensors," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 12, pp. 3574–3584, 2006.

[136] J. Karvonen, "Baltic sea ice concentration estimation using Sentinel-1 SAR and AMSR2 microwave radiometer data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 2871–2883, 2017.

[137] D. Malmgren-Hansen, L. T. Pedersen, A. A. Nielsen, M. B. Kreiner, R. Saldo, H. Skriver, J. Lavelle, J. Buus-Hinkler, and K. H. Krane, "A convolutional neural network architecture for Sentinel-1 and AMSR2 data fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 3, pp. 1890–1902, 2020.

[138] H. Lyu, W. Huang, and M. Mahdianpari, "A meta-analysis of sea ice monitoring using spaceborne polarimetric SAR: Advances in the last decade," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022.

[139] M. S. Mahmud, V. Nandan, S. Singha, S. E. Howell, T. Geldsetzer, J. Yackel, and B. Montpetit, "C-and L-band SAR signatures of Arctic sea ice during freeze-up," *Remote Sens. Environ.*, vol. 279, p. 113129, 2022.

[140] K. Kortum, S. Singha, and G. Spreen, "Robust multiseasonal ice classification from high-resolution X-band SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–12, 2022.

[141] L. Wang, A. Wong, D. A. Clausi, A. K. Scott, L. Xu, M. J. Shafiee, and F. Li, "Sea ice concentration estimation from satellite SAR imagery using convolutional neural network and stochastic fully connected conditional random field," in *CVPR 2015 Earthvision Workshop*. Citeseer, 2015.

[142] F. Li, D. A. Clausi, L. Wang, and L. Xu, "A semi-supervised approach for ice-water classification using dual-polarization SAR satellite imagery," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2015, pp. 28–35.

[143] J. Zhang, W. Zhang, Y. Hu, Q. Chu, and L. Liu, "An improved sea ice classification algorithm with Gaofen-3 dual-polarization SAR data based on deep convolutional neural networks," *Remote Sens.*, vol. 14, no. 4, p. 906, 2022.

[144] X. Chen, K. A. Scott, M. Jiang, Y. Fang, L. Xu, and D. A. Clausi, "Sea ice classification with dual-polarized SAR imagery: A hierarchical pipeline," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023, pp. 224–232.

[145] A. Kucik and A. Stokholm, "AI4SeaIce: selecting loss functions for automated SAR sea ice concentration charting," *Sci. Rep.*, vol. 13, no. 1, p. 5962, 2023.

[146] J. Karvonen, "Baltic sea ice concentration estimation from C-band dual-polarized SAR imagery by image segmentation and convolutional neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–11, 2021.

[147] Q. Yan and W. Huang, "Sea ice sensing from GNSS-R data using convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 10, pp. 1510–1514, 2018.

[148] L. Xu, D. A. Clausi, F. Li, and A. Wong, "Weakly supervised classification of remotely sensed imagery using label constraint and edge penalty," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 3, pp. 1424–1436, 2016.

[149] D. Hong, N. Yokoya, G.-S. Xia, J. Chanussot, and X. X. Zhu, "X-ModalNet: A semi-supervised deep cross-modal network for classification of remote sensing data," *ISPRS J. Photogramm. Remote Sens.*, vol. 167, pp. 12–23, 2020.

[150] Y. Wu, G. Mu, C. Qin, Q. Miao, W. Ma, and X. Zhang, "Semi-supervised hyperspectral image classification via spatial-regulated self-training," *Remote Sens.*, vol. 12, no. 1, p. 159, 2020.

[151] L. Yang, S. Yang, P. Jin, and R. Zhang, "Semi-supervised hyperspectral image classification using spatio-spectral laplacian support vector machine," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 3, pp. 651–655, 2013.

[152] X. Zhao, L. Liu, Y. Wang, L. Zhang, and A. Plaza, "DS4L: Deep semisupervised shared subspace learning for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

[153] Y. Shao, N. Sang, C. Gao, and L. Ma, "Spatial and class structure regularized sparse representation graph for semi-supervised hyperspectral image classification," *Pattern Recognit.*, vol. 81, pp. 81–94, 2018.

[154] Y. Zhao, F. Su, and F. Yan, "Novel semi-supervised hyperspectral image classification based on a superpixel graph and discrete potential method," *Remote Sens.*, vol. 12, no. 9, p. 1528, 2020.

[155] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[156] X. Zhang, S. Chen, P. Zhu, X. Tang, J. Feng, and L. Jiao, "Spatial pooling graph convolutional network for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022.

[157] Z. Wang, J. Li, T. Zhang, and S. Yuan, "Spectral–spatial discriminative broad graph convolution networks for hyperspectral image classification," *International Journal of Machine Learning and Cybernetics*, pp. 1–15, 2022.

[158] S. Kullback and R. A. Leibler, "On information and sufficiency," *The annals of mathematical statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[159] D. Hall, *Remote sensing of ice and snow.* Springer Science & Business Media, 2012.

[160] O. M. Johannessen, "Annex: Sar sea ice interpretation guide," in *Sea Ice in the Arctic.* Springer, 2020, pp. 507–573.

[161] R. Massom and D. Lubin, *Polar remote sensing.* Springer, 2006, vol. 2.

[162] M. Jiang, L. Xu, and D. A. Clausi, "Sea ice–water classification of RADARSAT-2 imagery based on residual neural networks (ResNet) with regional pooling," *Remote Sens.*, vol. 14, no. 13, p. 3025, 2022.

[163] M. Shokr and N. K. Sinha, *Sea ice: physics and remote sensing.* John Wiley & Sons, 2015.

[164] M. Dettwiler and A. (MDA), "Radarsat-2 product details," *Technical Report 1*, 2011.

[165] W.-I. Joint *et al.*, “Ice chart colour code standard, version 1.0, 2014.” World Meteorological Organization & Intergovernmental Oceanographic Commission, Tech. Rep., 2014.

[166] H. Liu, H. Guo, and G. Liu, “A two-scale method of sea ice classification using terrasar-x scansar data during early freeze-up,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021.

[167] N. Y. Zakhvatkina, V. Y. Alexandrov, O. M. Johannessen, S. Sandven, and I. Y. Frolov, “Classification of sea ice types in ENVISAT synthetic aperture radar images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 2587–2600, 2013.

[168] F. L. Hillebrand, I. D. de Carvalho Barreto, U. F. Bremer, J. Arigony-Neto, C. W. M. Júnior, J. C. Simões, C. N. da Rosa, and J. B. de Jesus, “Application of textural analysis to map the sea ice concentration with sentinel 1a in the western region of the antarctic peninsula,” *Polar Science*, vol. 29, p. 100719, 2021.

[169] deeplearningmisc, “deeplearningmisc,” *deeplearning*, vol. 8, no. 4, pp. 1601–1613, 2014.

[170] C. López-Martínez and X. Fabregas, “Polarimetric sar speckle noise model,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 10, pp. 2232–2242, 2003.

[171] R. Ressel, S. Singha, and S. Lehner, “Neural network based automatic sea ice classification for CL-pol RISAT-1 imagery,” in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2016, pp. 4835–4838.

[172] D. A. Clausi, “Comparison and fusion of co-occurrence, Gabor and MRF texture features for classification of SAR sea-ice imagery,” *Atmosphere-Ocean*, vol. 39, no. 3, pp. 183–194, 2001.

[173] M. Marbouti, O. Antropov, J. Praks, P. B. Eriksson, V. Arabzadeh, E. Rinne, and M. Leppäranta, “Tandem-x multiparametric data features in sea ice classification over the baltic sea,” *Geo-spatial information science*, vol. 24, no. 2, pp. 313–332, 2021.

[174] L. Wang, K. A. Scott, and D. A. Clausi, “Improved sea ice concentration estimation through fusing classified SAR imagery and AMSR-E data,” *Can. J. Remote Sens.*, vol. 42, no. 1, pp. 41–52, 2016.

[175] L. Wang, K. A. Scott, L. Xu, and D. A. Clausi, "Sea ice concentration estimation during melt from dual-pol SAR scenes using deep convolutional neural networks: A case study," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4524–4533, 2016.

[176] J. A. Karvonen, "Baltic sea ice SAR segmentation and classification using modified pulse-coupled neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 7, pp. 1566–1574, 2004.

[177] J. Karvonen, "Baltic sea ice concentration estimation based on C-band dual-polarized SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5558–5566, 2014.

[178] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep learning in remote sensing: a comprehensive review and list of resources," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 8–36, 2017.

[179] D. A. Clausi, "Texture segmentation of SAR sea ice imagery," Ph.D. dissertation, University of Waterloo, Waterloo, Ont, 1996.

[180] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM and MRF-based method for accurate classification of hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 4, pp. 736–740, 2010.

[181] D. A. Clausi and Y. Zhao, "Grey level co-occurrence integrated algorithm (GLCIA): a superior computational method to rapidly determine co-occurrence probability texture features," *Computers & Geosciences*, vol. 29, no. 7, pp. 837–850, 2003.

[182] D. A. Clausi and M. E. Jernigan, "A fast method to determine co-occurrence texture features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 1, pp. 298–300, 1998.

[183] D. A. Clausi and Y. Zhao, "Rapid extraction of image texture by co-occurrence using a hybrid data structure," *Computers & Geosciences*, vol. 28, no. 6, pp. 763–774, 2002.

[184] (2004) Ice chart colour code standard. [Online]. Available: https://en.vedur.is/media/hafis/frodleikur/ice-chart_colour-code-standard.pdf

[185] (2016) Manual of ice (MANICE), chapter 3: Observed ice charts. [Online]. Available: https://www.canada.ca/en/environment-climate-change/services/weather-manuals-documentation/manice-manual-of-ice/chapter-3.html

[186] (2016) Interpreting ice charts, chapter 1: The egg code. [Online]. Available: https://www.canada.ca/en/environment-climate-change/services/ice-forecasts-observations/publications/interpreting-charts/chapter-1.html

[187] P. Yu, A. Qin, and D. A. Clausi, "Feature extraction of dual-pol SAR imagery for sea ice image segmentation," *Can. J. Remote Sens.*, vol. 38, no. 3, pp. 352–366, 2012.

[188] Q. A. Holmes, D. R. Nuesch, and R. A. Shuchman, "Textural analysis and real-time classification of sea-ice types using digital SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, no. 2, pp. 113–120, 1984.

[189] R. W. Larson, J. D. Lyden, R. A. Shuchman, and R. T. Lowry, "Determination of backscatter characteristics of sea ice using synthetic aperture radar data," Environmental Research Inst of Michigan Ann Arbor Radar and Optics Div, Tech. Rep., 1981.

[190] Y. Hara, R. G. Atkins, R. T. Shin, J. A. Kong, S. H. Yueh, and R. Kwok, "Application of neural networks for sea ice classification in polarimetric sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 3, pp. 740–748, 1995.

[191] J. Fu, J. Liu, H. Tian, Y. Li, Y. Bao, Z. Fang, and H. Lu, "Dual attention network for scene segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 3146–3154.

[192] L.-K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 2, pp. 780–795, 1999.

[193] J.-W. Park, A. A. Korosov, M. Babiker, J.-S. Won, M. W. Hansen, and H.-C. Kim, "Classification of sea ice types in sentinel-1 synthetic aperture radar images," *The Cryosphere*, vol. 14, no. 8, pp. 2629–2645, 2020.

[194] M. Dabboor and T. Geldsetzer, "Towards sea ice classification using simulated radarsat constellation mission compact polarimetric sar imagery," *Remote Sens. Environ.*, vol. 140, pp. 189–195, 2014.

[195] W. Aldenhoff, C. Heuzé, and L. E. Eriksson, "Comparison of ice/water classification in Fram Strait from C-and L-band SAR imagery," *Annals of Glaciology*, vol. 59, no. 76pt2, pp. 112–123, 2018.

[196] Q. Zou, L. Ni, T. Zhang, and Q. Wang, "Deep learning based feature selection for remote sensing scene classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 11, pp. 2321–2325, 2015.

[197] P. Mohanaiah, P. Sathyanarayana, and L. GuruKumar, "Image texture feature extraction using GLCM approach," *International journal of scientific and research publications*, vol. 3, no. 5, pp. 1–5, 2013.

[198] M. Shokr, M. Dabboor, M. Lacelle, T. Zagon, and B. Deschamps, "Observations from C-band SAR fully polarimetric parameters of mobile sea ice based on radar scattering mechanisms to support operational sea ice monitoring," *Can. J. Remote Sens.*, pp. 1–17, 2021.

[199] K. Y. Vinnikov, A. Robock, R. J. Stouffer, J. E. Walsh, C. L. Parkinson, D. J. Cavalieri, J. F. Mitchell, D. Garrett, and V. F. Zakharov, "Global warming and northern hemisphere sea ice extent," *Science*, vol. 286, no. 5446, pp. 1934–1937, 1999.

[200] J. C. Comiso, D. J. Cavalieri, C. L. Parkinson, and P. Gloersen, "Passive microwave algorithms for sea ice concentration: A comparison of two techniques," *Remote Sens. Environ.*, vol. 60, no. 3, pp. 357–384, 1997.

[201] R. Jobanputra and D. A. Clausi, "Texture analysis using gaussian weighted grey level co-occurrence probabilities," in *First Canadian Conference on Computer and Robot Vision, 2004. Proceedings.* IEEE, 2004, pp. 51–57.

[202] W. G. Rees, *Remote sensing of snow and ice.* CRC press, 2005.

[203] M. Ghanbari, D. A. Clausi, L. Xu, and M. Jiang, "Unsupervised segmentation of multilook compact polarimetric sar data based on complex wishart distribution," in *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium.* IEEE, 2020, pp. 1456–1459.

[204] J. Chen, L. Jiao, X. Liu, L. Li, F. Liu, and S. Yang, "Automatic graph learning convolutional networks for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

[205] H. Zhao, F. Zhou, L. Bruzzone, R. Guan, and C. Yang, "Superpixel-level global and local similarity graph-based clustering for large hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

[206] H. Guo, Q. Shi, A. Marinoni, B. Du, and L. Zhang, “Deep building footprint update network: A semi-supervised method for updating existing building footprint from bi-temporal remote sensing images,” *Remote Sens. Environ.*, vol. 264, p. 112589, 2021.

[207] X. Sun, A. Shi, H. Huang, and H. Mayer, “BAS¡inline-formula¿¡tex-math notation=”latex”¿^{4}¡/tex-math¿¡/inline-formula¿Net: Boundary-aware semi-supervised semantic segmentation network for very high resolution remote sensing images,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 5398–5413, 2020.

[208] J. Hu, D. Hong, and X. X. Zhu, “Mima: Mapper-induced manifold alignment for semi-supervised fusion of optical image and polarimetric SAR data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 9025–9040, 2019.

[209] F. M. Riese, S. Keller, and S. Hinz, “Supervised and semi-supervised self-organizing maps for regression and classification focusing on hyperspectral data,” *Remote Sens.*, vol. 12, no. 1, p. 7, 2020.

[210] H. Zeng, Q. Liu, M. Zhang, X. Han, and Y. Wang, “Semi-supervised hyperspectral image classification with graph clustering convolutional networks,” *arXiv preprint arXiv:2012.10932*, 2020.

[211] Z. Zhang, “Semi-supervised hyperspectral image classification algorithm based on graph embedding and discriminative spatial information,” *Microprocessors and Microsystems*, vol. 75, p. 103070, 2020.

[212] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention.* Springer, 2015, pp. 234–241.

[213] W. Song, W. Gao, Q. He, A. Liotta, and W. Guo, “SI-STSAR-7: A large SAR images dataset with spatial and temporal information for classification of winter sea ice in Hudson Bay,” *Remote Sens.*, vol. 14, no. 1, p. 168, 2022.

[214] R. Shamshiri, E. Eide, and K. V. Høyland, “Spatio-temporal distribution of sea-ice thickness using a machine learning approach with Google Earth Engine and Sentinel-1 GRD data,” *Remote Sens. Environ.*, vol. 270, p. 112851, 2022.

[215] Y. Huang, Y. Ren, and X. Li, "Classifying sea ice types from SAR images using a U-Net-based deep learning model," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 3502–3505.

[216] Z. Sun, L. Sandoval, R. Crystal-Ornelas, S. M. Mousavi, J. Wang, C. Lin, N. Cristea, D. Tong, W. H. Carande, X. Ma *et al.*, "A review of earth artificial intelligence," *Computers & Geosciences*, p. 105034, 2022.

[217] J. A. Casey, S. E. Howell, A. Tivy, and C. Haas, "Separability of sea ice types from wide swath C-and L-band synthetic aperture radar imagery acquired during the melt season," *Remote Sens. Environ.*, vol. 174, pp. 314–328, 2016.

[218] W. Dierking and T. Busche, "Sea ice monitoring by l-band sar: An assessment based on literature and comparisons of JERS-1 and ERS-1 imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 4, pp. 957–970, 2006.

[219] S. E. Howell, A. S. Komarov, M. Dabboor, B. Montpetit, M. Brady, R. K. Scharien, M. S. Mahmud, V. Nandan, T. Geldsetzer, and J. J. Yackel, "Comparing L-and C-band synthetic aperture radar estimates of sea ice motion over different ice regimes," *Remote Sens. Environ.*, vol. 204, pp. 380–391, 2018.

[220] A. M. Johansson, C. Brekke, G. Spreen, and J. A. King, "X-, C-, and L-band SAR signatures of newly formed sea ice in Arctic leads during winter and spring," *Remote Sens. Environ.*, vol. 204, pp. 162–180, 2018.

[221] A. Gegiuc, J. Karvonen, J. Vainio, E. Rinne, R. Bednarik, and M. Mäkynen, "Visual interpretation of synthetic aperture radar sea ice imagery by expert and novice analysts: An eye tracking study," *The Cryosphere Discussions*, pp. 1–29, 2022.

[222] Z. Liang, X. Pang, Q. Ji, X. Zhao, G. Li, and Y. Chen, "An entropy-weighted network for polar sea ice open lead detection from Sentinel-1 SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.

[223] Y. Ren, X. Li, and W. Zhang, "A data-driven deep learning model for weekly sea ice concentration prediction of the Pan-Arctic during the melting season," *IEEE Transactions on Geoscience and Remote Sensing*, 2022.

[224] W. Guo, P. Itkin, S. Singha, A. Paul Doulgeris, M. Johansson, and G. Spreen, "Sea ice classification of TerraSAR-X ScanSAR images for the MOSAiC expedition incorporating per-class incidence angle dependency of image texture," *The Cryosphere Discussions*, pp. 1–29, 2022.

[225] S. Dominicus and A. K. Mishra, "Sea ice concentration estimation techniques using machine learning: An end-to-end workflow for estimating concentration maps from SAR images," *arXiv preprint arXiv:2205.01403*, 2022.

[226] X. Zhang, Y. Zhu, J. Zhang, Q. Wang, L. Shi, J. Meng, C. Fan, M. Liu, G. Liu, and M. Bao, "Assessment of arctic sea ice classification ability of chinese hy-2b dual-band radar altimeter during winter to early spring conditions," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 9855–9872, 2021.

[227] C. Tsatsoulis and R. Kwok, *Analysis of SAR data of the polar oceans: recent advances*. Springer Science & Business Media, 2012.

[228] M. Saarela and S. Jauhiainen, "Comparison of feature importance measures as explanations for classification models," *SN Applied Sciences*, vol. 3, no. 2, pp. 1–12, 2021.

[229] F. Liu, X. Qian, L. Jiao, X. Zhang, L. Li, and Y. Cui, "Contrastive learning-based dual dynamic gcn for sar image scene classification," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[230] X. Chen, K. A. Scott, L. Xu, M. Jiang, Y. Fang, and D. A. Clausi, "Uncertainty-incorporated ice and open water detection on dual-polarized SAR sea ice imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–13, 2023.

[231] A. S. Nagi, D. Kumar, D. Sola, and K. A. Scott, "RUF: Effective sea ice floe segmentation using end-to-end RES-UNET-CRF with dual loss," *Remote Sens.*, vol. 13, no. 13, p. 2460, 2021.

[232] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 14. Oakland, CA, USA, 1967, pp. 281–297.