Spatial Modeling of Compact Polarimetric Synthetic Aperture Radar Imagery

by

Saeid Taleghanidoozdoozan

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:	Avik Bhattacharya Professor, Centre of Studies in Resources Engineering, Indian Institute of Technology
Supervisor(s):	David A. Clausi Professor, Dept. of Systems Design Engineering, University of Waterloo
	Linlin Xu Research Assistant Professor, Dept. of Systems Design Engineering, University of Waterloo
Internal Member:	John Zelek Professor, Dept. of Systems Design Engineering, University of Waterloo
Internal-External Member:	Zhou Wang Professor, Dept. of Electrical and Computer Engineering University of Waterloo
	Grant Gunn Professor, Dept. 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 Contribution

The following three journal papers are used in this thesis. I was a co-author with a major contribution on the design, development, evaluation and writing of the material of papers.

Taleghanidoozdoozan, S., Xu, L., Clausi,D., A., (2022,published). "Oil Spill Candidate Detection Using a Conditional Random Field Model on Simulated Compact Polarimetric Imagery". *Canadian Journal of Remote Sensing*, vol. 48, no.3, pp.425-440.

Taleghanidoozdoozan, S., Xu, L., Clausi,D., A., (2023,Published). "Region-Based Sea Ice Mapping using Compact Polarimetric Synthetic Aperture Radar Imagery with Learned Features and Contextual Information". *Journal of Remote Sensing*, vol. 15, no.12.

Taleghanidoozdoozan, S., Xu, L., Clausi,D., A., (2023,Submitted). "Pyramid Fine- and Coarse-grained Attentions for Classifying Different Land Cover Types using Compact Polarimetric SAR Data". *IEEE Transactions on Geoscience and Remote Sensing.*

Abstract

The RADARSAT Constellation Mission (RCM) utilizes compact polarimetric (CP) mode to provide data with varying resolutions, supporting a wide range of applications including oil spill detection, sea ice mapping, and land cover analysis. However, the complexity and variability of CP data, influenced by factors such as weather conditions and satellite infrastructure, introduce signature ambiguity. This ambiguity poses challenges in accurate object classification, reducing discriminability and increasing uncertainty. To address these challenges, this thesis introduces tailored spatial models in CP SAR imagery through the utilization of machine learning techniques.

Firstly, to enhance oil spill monitoring, a novel conditional random field (CRF) is introduced. The CRF model leverages the statistical properties of CP SAR data and exploits similarities in labels and features among neighboring pixels to effectively model spatial interactions. By mitigating the impact of speckle noise and accurately distinguishing oil spill candidates from oil-free water, the CRF model achieves successful results even in scenarios where the availability of labeled samples is limited. This highlights the capability of CRF in handling situations with a scarcity of training data.

Secondly, to improve the accuracy of sea ice mapping, a region-based automated classification methodology is developed. This methodology incorporates learned features, spatial context, and statistical properties from various SAR modes, resulting in enhanced classification accuracy and improved algorithmic efficiency.

Thirdly, the presence of a high degree of heterogeneity in target distribution presents an additional challenge in land cover mapping tasks, further compounded by signature ambiguity. To address this, a novel transformer model is proposed. The transformer model incorporates both fine- and coarse-grained spatial dependencies between pixels and leverages different levels of features to enhance the accuracy of land cover type detection.

The proposed approaches have undergone extensive experimentation in various remote sensing tasks, validating their effectiveness. By introducing tailored spatial models and innovative algorithms, this thesis successfully addresses the inherent complexity and variability of CP data, thereby ensuring the accuracy and reliability of diverse applications in the field of remote sensing.

Acknowledgements

I would like to thank my supervisor Professor David Clausi for being a great mentor for me academically and personally. Thank you, Professor Clausi, for allowing me to grow as a researcher under your guidance. I look forward to many more collaborations in the future. I would also like to thank my co-supervisor, Research Assistant Professor Linlin Xu, for encouraging me to try all of the new ideas in my research in computer vision and remote sensing.

Many thanks to my Ph.D committee members Professor John Zelek, Professor Zhou Wang, Professor Grant Gunn for reviewing the thesis and providing their thoughtful comments on this thesis.

I would also like to thank my colleagues in the VIP lab and especially the remote sensing research group, including Mingzhe Jiang, Max lan Alexander Manning, Dr. Yuan Fang for their ideas and support. Special thanks to my colleague Dr. Mohsen Ghanbari for always being helpful in forming ideas, conducting experiments, and so on.

Thanks are also extended to the ice scientists in the Canadian Ice Service, including Melanie Lacelle, Marilee Pregitzer, Tom Zagon, and Benjamin Deschamps for providing labeled samples from the RADARSAT-2 scenes for this thesis.

Dedication

I dedicate my dissertation work to my family and many friends. A special feeling of gratitude to my loving parents, whose words of encouragement and push for tenacity ring in my ears.

I also dedicate this dissertation to my many friends who have supported me throughout the process. I will always appreciate all they have done, especially Dr.Mohsen Ghanbari for assisting me in developing my ideas, Max lan Alexander Manning for helping me in prepossessing data, and Mingzhe Jiang for being a constant source of encouragement during the challenges of my graduate program.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Compact Polarimetric Synthetic Aperture RADAR

Synthetic aperture radar (SAR) is an active remote sensing technique that provides images regardless of solar illumination and is relatively insensitive to weather conditions [1]. The RADARSAT Constellation Mission (RCM) is a new generation set of Earth observation SAR satellites, which was launched on June 12, 2019, and consists of three satellites [2]. The RCM data is used to support Canada's need for enhancing the operational use of SAR data for addressing three main key issues, namely, surveillance, disaster management, and environmental monitoring [2]. The RCM provides nearly daily coverage of Canada and it has dual-polarization (DP), wide-swath coverage compact polarimetry (CP), and quad-polarization (QP) modes [3]. In the RCM CP mode, the SAR system transmits a circularly-polarized radar signal and receives two coherent linear-polarized signals [4].

A CP SAR offers key advantages over QP SAR, including wide-swath coverage and reduced average transmitted power [5–7]. Moreover, the data acquired from CP SAR is expected to be of comparable quality to that obtained from a QP SAR [8]. On the other hand, DP SAR data provide broader area coverage than QP but offer limited detailed information about scattering objects, as they only produce intensity images. While both CP SAR and DP SAR share a wide swath width, CP SAR exhibits several benefits over DP SAR, including decreased sensitivity to noise and cross-channel errors, ease of implementation, self-calibration, and the ability to provide more comprehensive information about scattering objects [9].

RCM satellites offer CP data in various beam modes, ranging from high (3m) to low

(100m) resolutions, which can be used for various environmental monitoring purposes such as:

1. Oil spill detection. Detecting and efficiently mitigating the harmful impacts of oil spills, caused by accidents or illegal discharges, is critical for protecting marine ecosystems and human life [10, 11].

2. Sea ice mapping. The interpretation of ice types and the analysis of their properties have significant implications in various crucial applications, such as ship navigation, global climate monitoring, and animal migration forecasting [12].

3. Land cover classification. Accurate land cover classification is essential for urban planning, natural resource management, and environmental monitoring applications, as it provides valuable insights into changes occurring on the Earth's surface over time [13, 14].

Signature ambiguity, defined as the challenge of interpreting SAR data due to its inherent complexity and variability, poses significant challenges in classification methods. This highlights the necessity of incorporating spatial information to mitigate the impact of signature ambiguity and improve the accuracy and reliability of SAR data analysis and interpretation.

1.2 Spatial Information in CP SAR Imagery

Signature ambiguity is influenced by factors like weather conditions (e.g., wind speed and melting ponds on ice surface [15–17]) and satellite infrastructure (e.g., incidence angle effect [18–20] and speckle noise [21–23]). The presence of signature ambiguity significantly decreases discriminability and heightens uncertainty during data interpretation.

Spatial information allows for the consideration of neighboring pixels and their relationships, enabling the identification of coherent patterns and structures in the data [24]. By incorporating spatial information, the impact of ambiguities can be mitigated, leading to improved accuracy and reliability in SAR data analysis and interpretation [25]. This helps to enhance the discrimination capabilities of SAR classification algorithms.

There are various approaches to model spatial information in SAR images, including random field models, convolutional neural network (CNN), and transformer-based models. A conditional random field (CRF) model leverages statistical properties and similarities

in labels and features between pixels and their neighbors. CRF can reach good results in image classification tasks with limited labeled data [26]. Therefore, it can be used for tasks where annotated training data is scarce or expensive to obtain. On the other hand, CNNs extract spatial information by applying convolutional filters to capture local patterns and features, while pooling layers downsample the feature maps to retain salient information and capture abstract spatial patterns [27]. The ability of CNNs to effectively extract and learn fine-grained spatial features in SAR data [28] enhances their accuracy and effectiveness in accurately classifying various types of sea ice in SAR imagery. Another group of methods, transformers, extract spatial information by employing self-attention mechanisms to capture global dependencies, enabling the model to learn spatial relationships for effective information extraction [29]. By considering the high rate of spatial heterogeneity challenge in land cover classification tasks, modeling spatial information in different scales is important.

1.3 Motivation

The primary objective of this thesis is to develop advanced classification techniques specifically tailored for CP SAR imagery. By harnessing the power of spatial information within machine learning algorithms, the goal is to effectively address the complexity and variability encountered in CP SAR data and ultimately attain remarkably precise classification results. The motivations behind this research stem from several challenging aspects that warrant attention and exploration.

1. Previous approaches to oil spill detection using CP SAR have primarily relied on pixel-based features [30–36]. Nonetheless, the accuracy of single-pixel measurements is significantly impacted by signature ambiguity [37,38]. Furthermore, the statistical characteristics of CP SAR data have not been adequately considered in the design of classification methods. However, by incorporating the spatial information and statistical properties of CP SAR data, it becomes feasible to develop a CRF model that is specifically tailored to exploit the unique characteristics of CP SAR data without requiring extensive labeled samples.

2. The classification of sea ice using CP SAR data has been addressed in previous studies by incorporating various features and techniques, such as backscattering, polarimetric information, and texture features [4,39–46]. However, these methods rely on hand-crafted features, which necessitate domain expertise and time-consuming

feature engineering [47]. In contrast, deep learning methods offer a data-driven approach that can automatically learn features and spatial patterns from the data, eliminating the need for manual feature engineering [48]. Additionally, combining sea ice maps generated through pixel-level classification with segmented images has been shown to improve classification accuracy [43, 49]. However, previous CP sea ice classification approaches have overlooked the statistical properties of CP SAR data in the generation of regions. By considering these limitations, this is necessary to design a CP sea ice classification methodology that leverages deep learning techniques to automate feature extraction and incorporates the statistical properties of CP SAR data to generate accurate and reliable sea ice maps.

3. The classification of land cover using CP SAR data faces significant challenges, including signature ambiguity and spatial heterogeneity [50–52]. The limited research in this specific area highlights the need for improved methodologies [53–55]. Existing approaches fail to consider spatial information and learned features [53, 55], or lack contextual information [54]. To address these limitations, it is necessary to design a feature learning CP SAR classification that integrates spatial information in different levels and offers computational efficiency for large-scale CP scenes.

1.4 Objectives

Based on the signature ambiguity and spatial heterogeneity challenges and the limitations of the existing methods, this thesis aims to achieve the following three key objectives.

1. To improve the oil spill detection accuracy, this research aims to design a model based on a CRF which reduces the impacts of speckle noise and highlights the difference between oil spill candidates and oil-free water by effectively modeling the spatial contextual information and taking advantage of statistical characteristics of CP SAR data without requiring extensive training samples (See Chapter 3). Moreover, since solving the CRF is an ill-posed problem and optimization algorithm needs to be used to obtain the optimum result, this thesis assesses the performance of optimization algorithms so that the resulting CRF model can be tailored to the characteristics of the CP SAR data for dealing with the RCM CP SAR data in oil spill candidate detection applications.

2. To enhance the accuracy of sea ice mapping using CP SAR imagery, this thesis aims to develop a region-based automated sea ice classification methodology that

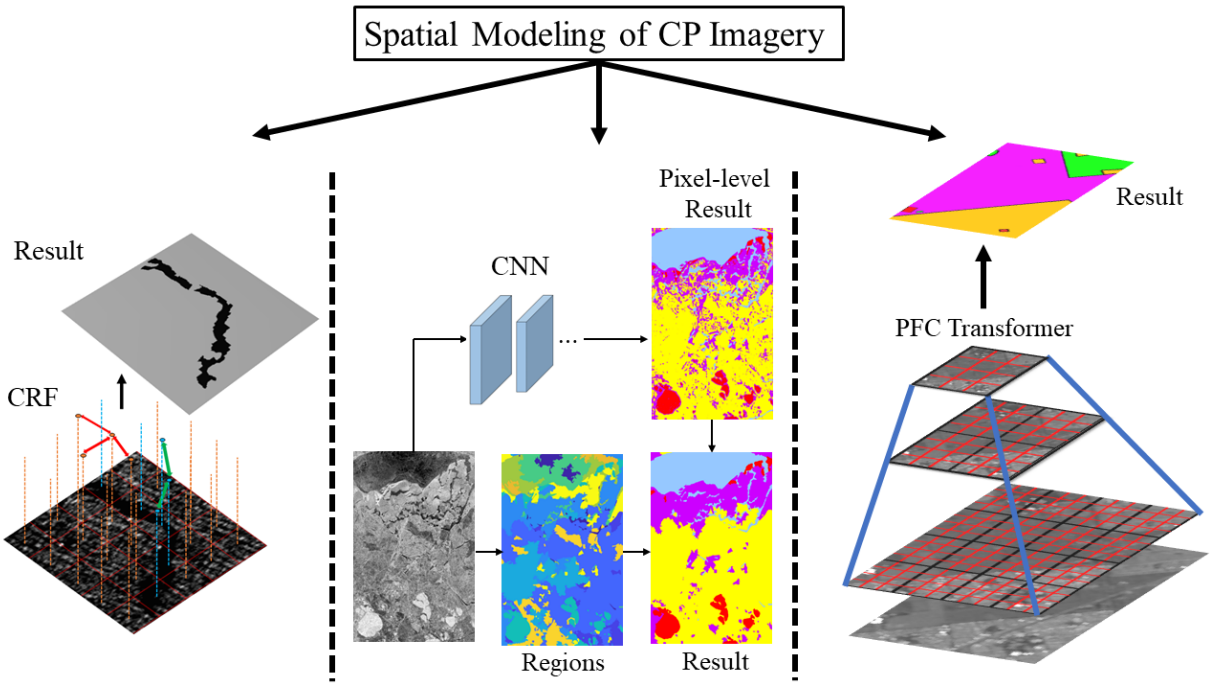


Figure 1.1: Different methods of modeling spatial information: CRF, Region-based CNN, and PFC Transformer.

incorporates learned features, spatial context, and statistical properties of CP SAR data (See Chapter4). In addition, the effectiveness of CP will be compared to DP, reconstructed QP (RQP), and QP modes for SAR sea ice classification, enabling the identification of the most suitable mode for improved operational algorithmic capability.

3. To improve the accuracy of the land cover mapping using CP SAR imagery, this thesis aims to design a novel approach called the pyramid fine- and coarse-grained self-attention transformer (PFC transformer) method (See Chapter 5). By leveraging fine- and coarse-grained spatial dependencies and integrating low- and high-level features, the PFC transformer method enables the extraction of detailed and comprehensive information from CP SAR imagery, ultimately leading to improved accuracy in land cover classification.

Fig 1.1 shows the three different methods of modeling spatial information.

1.5 Thesis Structure

In Chapters 3, 4, and 5, three distinct classification methods for CP SAR data are proposed. These chapters are based on three published/submitted journal manuscripts [56–58]. It is worth noting that the introduction to each chapter and the underlying theory may contain overlapping information.

The fundamentals of CP SAR data and their statistical properties are explained in Chapter 2. Chapter 3 has the objective of developing a CRF classification approach to identify oil spill candidates using CP SAR data. However, in the absence of CP SAR data for detecting oil spills, RADARSAT-2 scenes are utilized to simulate the corresponding CP SAR scenes. In Chapter 4, the proposed region-based sea ice classification method is explained. The structure of the PFC transformer method along with its advantage over the state-of-the-art attention-based methods are discussed in Chapter 5. Finally, a summary of contributions and conclusions as well as future works are described in Chapter 6.

Chapter 2

CP SAR Data

This chapter starts by introducing SAR data and statistical properties of CP SAR data. From there, the advantages of CP SAR imagery system over QP and DP ones are discussed. Finally, the theory of reconstructing QP SAR data from CP is explained.

2.1 Synthetic Aperture Radar

Remote sensing systems can be categorized in two general groups: a) passive and 2) active. A passive remote sensing system records electromagnetic energy emitted (e.g., thermal infrared radiation) or reflected (e.g., blue, green, red, and near-infrared light) from Earth's surface [1]. In contrast to passive, active remote sensing systems rely neither on the Sun's electromagnetic energy nor the Earth's thermal properties. In an active remote sensing system, electromagnetic energy is transmitted from the sensor toward the terrain. Interacting the transmitted wave with the terrain produces a backscatter of energy recorded by the sensors [1]. The following active remote sensing systems are the most widely used: RADAR, LIDAR, and SONAR. For Earth-resource observations, SAR remote sensing is most commonly used [1].

SAR satellites of the first generation, such as SEASAT, transmit and receive intensity data over a single channel in HH (horizontal linear transmission and reception), VV (vertical linear transmission and reception), HV (horizontal linear transmission and vertical linear reception), and VH (vertical linear transmission and horizontal linear reception) [59]. In single-polarized SAR images, intensity values and texture features are mainly used to classify SAR scenes. Providing less information is the main limitation of single-polarized

SAR systems, particularly in dealing with ice-covered water bodies [59]. DP SAR systems such as RADARSAT-2 provide two intensity channels (HH/HV, VV/VH, or HH/VV). Canadian Ice Service (CIS) experts mainly use DP RADARSAT-2 data to generate sea ice maps [43]. The main advantage of DP SAR imagery is its wide swath width. However, the phase information between channels is not measured in this SAR system, resulting in incomplete information about objects [60]. In contrast to DP, QP SAR systems obtain complete polarimetric information about objects by saving the phase information between channels [60]. CP SAR systems provide more information than DP and less than QP. In other words, CP SAR systems take advantage of DP and QP SAR data by providing wide swath width and preserving phase information between channels [8].

2.2 Introduction to SAR Polarimetry

The electric field orientation of an electromagnetic (EM) wave in space is defined as polarization [60]. Polarization of an electric field explains the behavior of a field and can be characterized by using the Jones vector $\underline{E}$ [60]:

$$\underline{E} = \begin{bmatrix} E_{0x}e^{j\delta_x} \\ E_{0y}e^{j\delta_y} \end{bmatrix} \tag{2.1}$$

The Jones vector characterizes a wave by decomposing it into two orthogonal waves with (E_{0x}, E_{0y}) amplitudes and (δ_x, δ_y) phases. Different values of phases and amplitudes define different polarizations. For example, linear and circular polarizations are defined by $(\delta_x - \delta_y = 0)$ and $\left(\delta_x - \delta_y = \frac{\pi}{2}, E_{0x} = E_{0y}\right)$, respectively.

Another way to describe polarization of an EM wave is using the polarization ellipse including three parameters, namely, the ellipse amplitude A, the orientation angle ϕ, and the ellipticity angle τ (see Fig 2.1). Using polarization ellipse parameters, a Jones vector can be formulated as follows [60]:

$$\underline{E} = Ae^{j\alpha} \begin{bmatrix} \cos\phi\cos\tau - j\sin\phi\sin\tau \\ \sin\phi\cos\tau + j\cos\phi\sin\tau \end{bmatrix} \tag{2.2}$$

where α is an absolute phase term and A$= \sqrt{E_{0x}^2 + E_{0y}^2}$. By using (2.2), the unit Jones vector associated with a canonical polarization can be calculated. Table 2.1 shows the unit Jones vectors of different canonical polarizations.

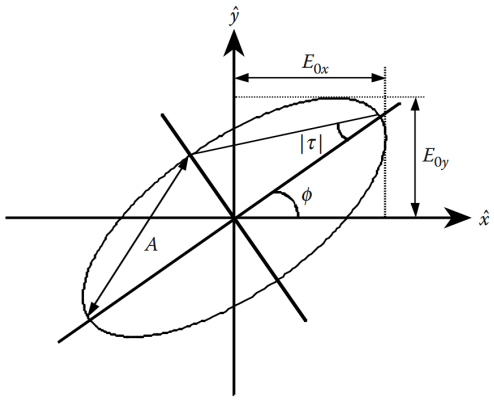


Figure 2.1: Polarization ellipse [60]

In the case of a SAR system, the vector $\underline{E}$ of the backscattered field is calculated as [60]:

$$\underline{E} = \mathbf{S}\underline{\hat{u}}_t = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \underline{\hat{u}}_t \tag{2.3}$$

where $\underline{\hat{u}}_t$ is a unit Jones vector associated with a canonical polarization and $\mathbf{S}$ is scattering matrix. S_{ij} is a complex number which i and j show transmitted and received polarization basis, respectively [60]. The diagonal elements of scattering matrix are known as co-polarized coefficients while the off-diagonal ones are called cross-polarized coefficients. When the same antenna is used for transmitting and receiving a wave, the reciprocal theorem requires that $S_{HV} = S_{VH}$ [60].

A QP SAR system provides a complete scattering matrix; however, it requires a higher pulse repetition frequency than a DP SAR to transmit and receive in two orthogonal polarizations. In a QP system, the pulse repetition frequency is twice that of a single or DP SAR for a given coverage area [8]. However, the swath width of QP SAR imagery is limited, making it not appropriate for ocean application, particularly generating sea ice maps [61]. Therefore, using CP mode is more appropriate for ocean applications.

2.3 Compact Polarimetry

The architectures of CP SAR are classified into (i) $\pi/4$, (ii) the circularly-circularly (CC), and (iii) the hybrid-polarity (HP) modes [31]. In $\pi/4$ mode, a linear polarized field with a 45° inclination is transmitted and two coherent linear polarizations are received. In $\pi/4$

Table 2.1: The Jones vector of canonical polarizations

Polarization state	Unit Jones vector $\underline{\hat{u}}_t$	Orientation angle ϕ	Ellipticity angle τ
Horizontal	$\begin{bmatrix}1\\0\end{bmatrix}$	0	0
Vertical	$\begin{bmatrix}0\\1\end{bmatrix}$	$\frac{\pi}{2}$	0
Linear +45°	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$	$\frac{\pi}{4}$	0
Linear -45°	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$	$-\frac{\pi}{4}$	0
Left circular	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\j\end{bmatrix}$	$[-\frac{\pi}{2}...\frac{\pi}{2}]$	$\frac{\pi}{4}$
Right circular	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-j\end{bmatrix}$	$[-\frac{\pi}{2}...\frac{\pi}{2}]$	$-\frac{\pi}{4}$

mode, the orientation invariance for double-bounce responses is not be ensured [9]. To overcome this issue, a circular polarization must be transmitted. In CC mode, a circularly polarized filed is transmitted and the response of the target is received in the same polarization. In terms of hardware, the implementation of CC architecture is difficult [31]. HP mode proposed by Raney [9] can be an alternative of CC mode. In HP mode, the recording system transmits a circular polarization field and receives coherently both horizontal and vertical polarizations. Since the coherency matrix is independent of the receiver's polarization basis, then in HP architecture not only all the advantages of CC one are maintained,

but also the implementation of hardware becomes simpler [9]. Moreover, HP mode is rotationally robust with respect to the shape of the observed scene [31].

Considering an HP mode in which a right circular polarized wave is transmitted and coherent dual linear polarizations are received, the unit Jones vector is $[\frac{1}{\sqrt{2}}, \frac{-j}{\sqrt{2}}]^T$ (see Table 2.1). Therefore, $\underline{E}_{HP}$ is calculated as follows:

$$\underline{E}_{HP} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-j}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} - jS_{HV} \\ S_{HV} - jS_{VV} \end{bmatrix} \tag{2.4}$$

Because of some reasons such as increasing the information content of polarimetric SAR data and reducing the impacts of speckle noise, the coherency matrix is used instead of scattering matrix [62]. The coherency matrix of a CP SAR data is a 2×2 positive semi-definite Hermitian matrix which is defined as [63]:

$$\mathbf{J}_{CP} = \frac{1}{L} \sum_{i=1}^{L} \underline{E} \cdot \underline{E}^{*^T} = \begin{bmatrix} J^{11} & J^{12} \\ J^{21} & J^{22} \end{bmatrix} \tag{2.5}$$

where L is the number of looks for averaging. T and $*$ stand for the transpose operator and complex conjugate, respectively. According to (2.4) and (2.5), the coherency matrix of $\mathbf{J}_{HP}$ can be given by [64]

$$\begin{aligned} \mathbf{J}_{HP} = \frac{1}{2} \begin{bmatrix} \langle \mid S_{HH} \mid^2 \rangle & j\langle S_{HH} S_{VV}^* \rangle \\ -j\langle S_{VV} S_{HH}^* \rangle & \langle \mid S_{VV} \mid^2 \rangle \end{bmatrix} + \frac{\langle \mid S_{HV} \mid^2 \rangle}{2} \begin{bmatrix} 1 & -j \\ j & 1 \end{bmatrix} + \\ \frac{1}{2} \begin{bmatrix} -2Im\left(\langle S_{HH} S_{HV}^* \rangle\right) & \langle S_{HH} S_{HV}^* \rangle + \langle S_{VV}^* S_{HV} \rangle \\ \langle S_{HH}^* S_{HV} \rangle + \langle S_{VV} S_{HV}^* \rangle & 2Im\left(\langle S_{VV} S_{HV}^* \rangle\right) \end{bmatrix} \end{aligned} \tag{2.6}$$

where Im indicates the imaginary part of a complex number.

2.3.1 Statistical Characteristics of CP SAR Data

Statistical properties of SAR data are considered in several tasks, parametric classification models for instance. The one look amplitude SAR image follows a Rayleigh probability distribution while the intensity image has a negative exponential distribution [60]. In many studies, it is assumed that a SAR image follows the mixture Gaussian distribution [65]. Assuming random variable y represents a pixel value in a SAR image, for example intensity

value of a pixel in an HH image of RADARSAT-2, the probability density function (PDF) of y is:

$$p\left(y|\mu_i,\sigma_i\right)=\frac{1}{\sigma_i\sqrt{2\pi}}exp\left(\frac{-\left(y-\mu_i\right)^2}{2\sigma_i^2}\right). \tag{2.7}$$

μ_i and σ_i are the mean and standard deviation of the class i, respectively.

To reduce speckle noise and compress SAR data, multi-look processing is performed by averaging several single-look images [60, 66]. It has been shown that an L-look SAR intensity image has a Chi-square distribution with 2L degrees of freedom [60,67]. Therefore, PDF of a pixel in an L-look SAR image is described by [66, 68, 69]:

$$p\left(y|\mu_i,L\right)=\frac{1}{\Gamma\left(L\right)}\left(\frac{L}{\mu_i}\right)^L y^{L-1}exp\left(-\frac{Ly}{\mu_i}\right) \tag{2.8}$$

where $\Gamma\left(.\right)$ is gamma function and μ_i stand for the mean of the i^{th} class [70]. Moreover, L can be estimated as the ratio of the mean and the variance of the intensity [66, 71].

The coherency matrix for SAR data includes complex numbers and has a complex Wishart distribution [72]. The kth pixel in CP SAR imagery is represented by a complex matrix $\mathbf{A}_k=L\mathbf{J}_k$. The PDF of the matrix $\mathbf{A}_k$ is given by [73],

$$P_{\mathbf{A}_k}^{(L)}\left(\mathbf{A}_k|J_i\right)=\frac{|\ \mathbf{A}_k\ |^{L-q}\ exp\left[-Tr\left(\mathbf{J}_i^{-1}\mathbf{A}_k\right)\right]}{K\left(L,q\right)\ |\ \mathbf{J}_i\ |^L} \tag{2.9}$$

where $\mathbf{J}_i$ is the coherency matrix of the i^{th}class . Tr(.) and $|\ .\ |$ stand for trace of (.) and the determinant operator, respectively. Also,

$$K\left(L,q\right)=\pi^{\frac{1}{2}q(q-1)}\Gamma\left(L\right),\cdots,\Gamma\left(L-q+1\right) \tag{2.10}$$

where the parameter q displays the dimension of the scattering vector of a SAR data which for CP data is 2.

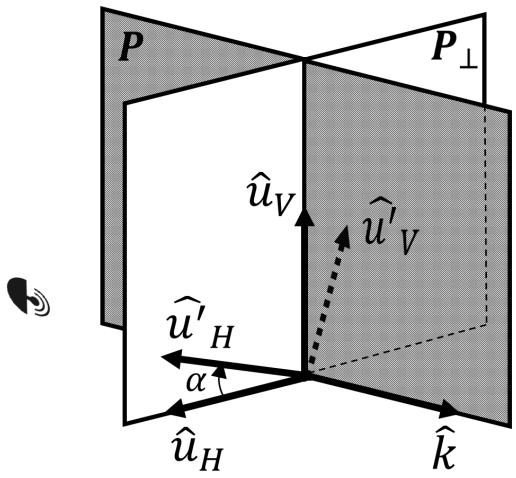


Figure 2.2: Traverse $P_\perp$ and incident P planes involved in reflection symmetry (This figure is based on Fig. 1 by Souyris *et al.* [74].)

2.4 Reconstructing QP SAR Data from CP

Assuming reciprocity in the monostatic radar case, QP SAR data provides a 3×3 covariance matrix as follows [73]:

$$\mathbf{C}_{QP} = \begin{bmatrix} \langle |S_{HH}|^2 \rangle & \sqrt{2}\,\langle S_{HH} S_{HV}^* \rangle & \langle S_{HH} S_{VV}^* \rangle \\ \sqrt{2}\,\langle S_{HV} S_{HH}^* \rangle & 2\,\langle |S_{HV}|^2 \rangle & \sqrt{2}\,\langle S_{HV} S_{VV}^* \rangle \\ \langle S_{VV} S_{HH}^* \rangle & \sqrt{2}\,\langle S_{VV} S_{HV}^* \rangle & \langle |S_{VV}|^2 \rangle \end{bmatrix} = \begin{bmatrix} C^{11} & C^{12} & C^{13} \\ C^{21} & C^{22} & C^{23} \\ C^{31} & C^{32} & C^{33} \end{bmatrix}. \quad (2.11)$$

Based on the CP SAR data, the QP SAR data can be reconstructed to be used as an alternative representation of the CP SAR data to utilize the full polarimetric SAR classification methods [43] and provide access to known QP parameters whose usage and interpretation are known [4]. To reconstruct the 3×3 coherency matrix of QP SAR data by using the 2×2 CP coherency matrix, according to (2.5) and (2.11), there are four known equations (two real J^{11} and J^{22} channels, and a complex J^{12} channel) and nine unknown variables (three real C^{11}, C^{22}, and C^{33} channels and three complex C^{12}, C^{13}, and C^{23} ones). It is necessary to make some assumptions to solve this underdetermined system of equations [74, 75].

Souyris *et al.* [74] proposed the first QP reconstruction method based on reflection and rotation symmetry assumptions [74, 76] and Nord *et al.* [75] modified it. Assume that ($\hat{\mathbf{u}}_H \perp P$) and ($\hat{\mathbf{u}}_V \parallel P$) where plane P contains the incident wave $\hat{k}$ and $P_\perp$ is the traverse

plane normal to the incident wave (see Fig. 2.2). When a target is symmetric about the plane P, reflection symmetry is formed. The reflection symmetry can be observed in various geophysical media, such as ocean water, forests, snow, and sea ice [74]. The reflection symmetry results in an approximate zero correlation between the co- and cross-pol coefficients, $\langle S_{HH}S_{HV}^{*}\rangle = \langle S_{VV}S_{HV}^{*}\rangle = 0$. Consequently, the covariance matrix of reconstructed QP (RQP) is formed as [76]:

$$\mathbf{C}_{RQP} = \begin{bmatrix} \langle|S_{HH}|^2\rangle & 0 & \langle S_{HH}S_{VV}^{*}\rangle \\ 0 & 2\langle|S_{HV}|^2\rangle & 0 \\ \langle S_{VV}S_{HH}^{*}\rangle & 0 & \langle|S_{VV}|^2\rangle \end{bmatrix} \tag{2.12}$$

and $\mathbf{J}_{HP}$ with the reflection symmetry is given as:

$$\mathbf{J}_{HP} = \frac{1}{2}\begin{bmatrix} \langle|S_{HH}|^2\rangle + \langle|S_{HV}|^2\rangle & j\left(\langle S_{HH}S_{VV}^{*}\rangle - \langle|S_{HV}|^2\rangle\right) \\ -j\left(\langle S_{VV}S_{HH}^{*}\rangle - \langle|S_{HV}|^2\rangle\right) & \langle|S_{VV}|^2\rangle + \langle|S_{HV}|^2\rangle \end{bmatrix} \tag{2.13}$$

Rotation symmetry is another assumed symmetry. The covariance matrix coefficients of a target with rotation symmetry around the incident wave are invariant to the rotation of the orthogonal basis $(\hat{\mathbf{u}}_H, \hat{\mathbf{u}}_V)$ by any arbitrary angle α (see Fig. 2.2). Under rotation symmetry assumption, Nord *et al.* [75] established a general relationship between the linear HH-VV coherence and the cross-polarization ratio

$$\frac{\langle|S_{HV}|^2\rangle}{\langle|S_{HH}|^2\rangle + \langle|S_{VV}|^2\rangle} = \frac{1 - |\rho_{HHVV}|}{N} \tag{2.14}$$

where ρ_{HHVV} and N are given by

$$\rho_{HHVV} = \frac{\langle|S_{HV}|^2\rangle - 2jJ^{12}}{\sqrt{(2J^{11} - \langle|S_{HV}|^2\rangle)(2J^{22} - \langle|S_{HV}|^2\rangle)}} \tag{2.15}$$

$$N = \frac{\langle|S_{HH} - S_{VV}|^2\rangle}{\langle|S_{HV}|^2\rangle} \tag{2.16}$$

The reconstruction algorithm of $\mathbf{C}_{RQP}$ consists of an iterative process that involves alternating between (2.15) and (2.14). The iteration begins with setting N to 4. After computing $\mathbf{C}_{RQP}$, the value of N is estimated using (2.16) and then utilized for calculating the new $\mathbf{C}_{RQP}$ (see [74, 75] for more details).

Chapter 3

Oil Spill Candidate Detection Using a Conditional Random Field Model on Simulated Compact Polarimetric Imagery

Although the CP SAR mode of the RCM offers new opportunities for oil spill candidate detection, there has not been an efficient machine learning model explicitly designed to utilize this new CP SAR data for improved detection. This chapter presents a conditional random field model based on the Wishart mixture model (CRF-WMM) to detect oil spill candidates in CP SAR imagery. First, a "Wishart mixture model" (WMM) is designed as the unary potential in the CRF-WMM to address the class-dependent information of oil spill candidates and oil-free water. Second, we introduce a new similarity measure based on CP statistics designed as a pairwise potential in the CRF-WMM model so that pixels with strong spatial connections have the same class label. Finally, we investigate three different optimization approaches to solve the resulting maximum a posterior (MAP) problem, namely iterated conditional modes (ICM) [77], simulated annealing (SA) [78], and graph cuts (GC) [79]. The results show that our proposed CRF-WMM model can delineate oil spill candidates better than the traditional CRF approaches, and that the GC algorithm provides the best optimization. This chapter is based on a published paper [56].

3.1 Introduction

Protecting the environment is an important subject that has gained significant attention across the world. One of the most damaging events to the marine ecosystem is an oil spill which is usually caused by oil tanker accidents, illegal oily discharges from tank cleaning, or oil pipeline breakages [80]. As a result, detecting and cleaning up oil spills play an essential role in the lives of humans and marine life. The first step to efficiently reduce the destructive effects of oil spills on the environment is to detect oil spills accurately.

An appropriate choice to detect marine oil spills is to use SAR systems because they can capture images of large areas, regardless of image acquisition time and weather conditions. Oil spill detection can be done manually by trained operators. In visual inspection methods, an expert has to examine the entire scene and identify oil spills [81]. Although a trained operator can identify oil spill candidates in a SAR image with sufficient certainty, processing SAR scenes by human analysts is difficult and time-consuming. Outputs generated from the visual inspection can vary since these are dependent on the knowledge and experience of operators. Therefore, designing automatic models to use SAR data for oil spill detection is essential for accurately detecting oil spills. In general, the process of detecting oil spills from SAR imagery can be divided into three steps: 1) detection of candidate oil spill spots, 2) extraction of features from the candidate spots, and 3) classification of candidate oil spots [82]. The first step is the most important part because if an oil spill is not detected, it is impossible to retrieve it in the rest steps.

A number of natural phenomena can result in false oil spill detections in SAR imagery such as marine organisms, shear zones, natural low wind zones (wind speed < 3 m/s), internal waves, rain cells, grease ice, and microconvective cells, which are known as oil spill look-alikes [83, 84]. The normalized radar cross section (NRCS) of look-alikes and oil spills are both low and similar making it a major problem in oil spill detection. The coherent interference of backscattered waves from many randomly distributed scatterers within a resolution cell causes bright and dark pixels in a radar image [1, 85]. This grainy salt-and-pepper noise is called speckle noise. An essential challenge in SAR oil spill candidate detection is caused by the speckle noise that creates low feature space separability between the oil spill and oil-free water classes. Generally, to reduce the effects of speckle noise, incoherent averaging using multiple looks or adaptive/non-adaptive signal processing filters are utilized [86].

Only a few published papers exist on the topic of oil spill detection using CP SAR data [30, 31, 34, 87]. Previous studies have not used the statistical distribution of CP SAR data in oil spill detection applications. Therefore, one of the main contributions of this

chapter lies in incorporating statistical properties of CP SAR data to identify oil spill candidates. Context is an important information source in analyzing SAR data. In this chapter, the spatial context information in SAR imagery is used to improve class separability. We introduce a model based on a CRF which reduces the impacts of speckle noise and highlights the difference between oil spills and oil-free water by effectively modeling the spatial information. The CRF consists of the unary and pairwise potentials. The unary potential utilizes the statistical properties of individual pixels, while the pairwise potential represents relationships between neighboring pixels [88,89]. Nevertheless, a CRF is traditionally built upon SAR intensity images, and we are not aware of a CRF model designed to address the statistical properties of the complex CP SAR data. The second aim of this chapter is to introduce a similarity measure based on CP SAR data to better model spatial information. Since solving the CRF is an ill-posed problem, an optimization algorithm needs to be used to obtain the optimum result. Therefore, the third aim of this chapter is to assess the performance of optimization algorithms so that the resulting CRF model can be tailored to the characteristics of the CP SAR data for dealing with the future RCM CP SAR data in oil spill candidate detection applications.

In summary, the contributions of this chapter lie in the following aspects. First, a Wishart mixture model (WMM) based on the complex Wishart distribution is designed to implement the CRF-WMM's unary potential to accommodate the discriminative class statistics in the CP SAR data. Second, the pairwise potential in CRF-WMM is implemented to leverage spatial information. Finally, we identify the best approach for solving the resulting CRF-WMM model of the three popular optimization approaches.

A literature review on the related workers are presented in Section 3.2. In Section 3.3, the details of the proposed CRF is described. Then, after introducing the study area in Section 3.4, the results are presented in Section 3.5. The last section concludes the chapter.

3.2 Related Work

Early research on SAR oil spill detection was conducted using single-polarization SAR data, generally using the VV (vertical transmit, vertical receive) polarization [90]. Elachi [91] is recognized as the first to use SAR data to monitor oil spills using data from the Seasat satellite. The enhanced capabilities of the second generation of satellite SAR sensors such as ENVISAT, RADARSAT-1, and ERS-2 attracted further attention for oil spill monitoring [92–94]. Xu *et al.* [95] performed a comprehensive study of different classification techniques to detect oil spills using ninety-three RADARSAT-1 ScanSAR Narrow Beam

images. They utilized fifteen features and different classification methods to distinguish oil spills from look-alike phenomena.

The third generation of SAR sensors, such as the Canadian RADARSAT-2 and German TerraSAR-X, offer QP SAR modes, higher spatial resolution and shorter revisit times than previous SAR missions [96, 97]. With the availability of QP SAR data, researchers have also explored QP SAR data for oil spill selection [31, 98, 99]. Minchew *et al.* [100] investigated $H/A/\sqrt{\alpha}$ eigenvector decomposition parameters extracted from QP UAVSAR (uninhabited aerial vehicle SAR) data to analyze the backscattering of the Deepwater Horizon (DWH) oil spill and determined that the major eigenvalue of the coherency matrix was the most promising indicator for oil slick detection. Genovez *et al.* [101] proposed a multisource approach to utilize optical, single-channel SAR, and QP SAR data to distinguish oil from water and classify oil into two thick and thin layers. Espeseth *et al.* [102] used a series of short time revisit SAR images to identify areas with relatively thick oil slicks. Their results showed that multiple SAR images with short repeat times could provide new information to identify short term oil slick drifts, which is important for clean up efforts. Skrunes *et al.* [103] investigated the discrimination potential of the eight well-known multipolarization features by measuring between-region contrast and within-region variance and concluded that the pair of geometric intensity and the real part of the copolarization cross product features could be used to determine the most promising results.

Several approaches have been proposed for oil spill candidate detection by taking advantage of spatial information. One method is to employ a graphical model. Xu *et al.* [104] introduced a stochastic fully connected CRF based on a Gaussian mixture model (GMM) to detect dark spots on RADARSAT-1 images, demonstrating that incorporating spatial information can improve results. Morales *et al.* [105] utilized a hierarchical Markov random field (MRF) to segment SAR images into oil classes (denser and thinner) and sea water. Pelizzari and Bioucas-Dias [106] used an MRF based on graph cuts to detect oil spills in SAR intensity images. Martinis [107] used a hybrid Markov image model by integrating scale-dependent and spatial information into the labeling process for near real-time oil spill detections in high-resolution TerraSARX ScanSAR data. Parmiggiani *et al.* [108] used a threshold GMM-MRF model to segment oil spills in a SAR image. Other mathematical tools which utilize spatial relationships among pixels are convolution neural networks (CNNs). De Laurentiis *et al.* [83] utilized a CNN to separate mineral films from biogenic slicks and oil-free sea surface. Guo *et al.* [109] employed a CNN to discriminate oil spills from look-alikes, while Yaohua and Xudong [110] used a dense connected CNN for the same task.

There are a limited number of published papers on oil spill detection using CP SAR data. Salberg *et al.* [30] proposed a coherence measure, which relies on the Bragg scattering

assumption, to detect oil spills in simulated HP SAR data obtained from RADARSAT-2 data. They compared the performance of their proposed measure with the degree of polarization (DoP), the conformity coefficient, and the correlation coefficient. According to their results, the coherence measure suppresses some look-alikes caused by low wind. Nunziata *et al.* [31] investigated features extracted from HP SAR data based on sea surface scattering with or without oil spills. Moreover, they proposed a new measure called the standard deviation of the phase difference (σ), where the value of σ for oil-covered surfaces is larger than that for slick-free regions. Li *et al.* [32] proposed a new method to retrieve the mixture ratio of oil and water in CP SAR data by using the diagonal elements of the covariance matrix of CP SAR data. In [33], an iterative reconstruction QP SAR image method using CP SAR data based on the different statistical behavior between oil spills and open sea surface was proposed. Shirvany *et al.* [34] investigated the potential use of DoP calculated for CP and DP SAR data to detect ships and oil spills. Their experiment showed that using the HP CP mode and DP (HH , VV) modes result in better detection performance. Espeseth *et al.* [35] compared the performance of a set of features extracted from QP UAVSAR and simulated HP SAR data under a high wind situation, and concluded that the capability of the HP data to distinguish different slicks from each other as well as from oil-free water is comparable with that of the QP data. Collins *et al.* [90] investigated the potential of the RQP SAR data based on CP SAR data to create the oil/water mixing index called Mdex [111]. According to their results, RQP SAR and QP SAR data resulted in similar Mdex maps. Chaudhary and Kumar [112] investigated the potential of using decomposition parameters extracted from QP UAVSAR (uninhabited aerial vehicle SAR) and CP RISAT-1 SAR images for detecting oil slicks. Here, using Van Zyl parameters [113] achieved the best results for the UAVSAR dataset, and utilizing compact-pol decomposition parameters [114] achieved the best results for the RISAT-1 dataset. A study by Chaudhary and Kumar [115] investigated the capability of using features measured from QP and simulated HP imagery to distinguish oil spills from oil-free water. Classification using HP features achieved an accuracy of more than 98%. Li *et al.* [36] analyzed the polarimetric properties of oil-covered ocean surface water in CP SAR data using the polarimetric degree m and the Poincare χ parameters [116]. They concluded that the sign of χ is opposite for the oil spills from oil-free water, and compared to oil-free water, oil spills reduce the value of m.

Most of the methods described above have been applied to detect oil spills by utilizing pixel-based features. However, single-pixel measurements are strongly affected by speckle noise resulting in noisy outputs [37,38]. In this chapter, we consider spatial information to reduce speckle effects and the impact of intra-class variations to increase class separability for the oil spill detection problem. Moreover, in contrast to previous studies such as [104,

117], we investigate the use of a more appropriate statistical model to include in a CRF model that is better able to more accurately detect oil spill candidates in CP SAR data.

3.3 Methodology

3.3.1 CRF-WMM Framework

Let S represent the discrete two-dimensional rectangular lattice and $i \in S$ be a site in the lattice (i.e., a pixel in the image). A CP SAR image is represented by $Y = \{y_i | i \in S\}$ where Y is a random field on S. Moreover, $X = \{x_i | i \in S\}$ is a set of binary-valued random variables which form a random field on S. For oil spill candidates x_i=1, oil-free water x_i=0. Based on this formulation, the task of oil spill candidate detection from CP SAR imagery aims to estimate X based on Y, which is achieved using a novel CRF-WMM model that addresses both the backscattered information and the spatial context information in the CP image.

The proposed CRF-WMM model addresses the posterior probability distribution of the label map X given the CP SAR imagery observation Y by [104]:

$$p(X|Y) = \frac{1}{Z(X)} exp\{-\sum_{i \in S} \psi_u(x_i, y_i) - \beta \sum_{i \in S} \sum_{j \in N_i} \psi_p(x_i, x_j, y_i, y_j)\} \tag{3.1}$$

in which $Z(X)$ is a normalization factor called the partition function. ψ_u and ψ_p are the unary potential and the pairwise potential, respectively, and β dictates the relative weight of the two potentials. N_i denotes a set of 4 neighbors of pixel i. Given this formulation, the estimation of X is achieved by maximizing $p(X|Y)$ under the maximum a posterior (MAP) framework. The CRF-WMM model is solved by addressing the following MAP problem:

$$\hat{X} = arg \max_X (p(X|Y)) \tag{3.2}$$

To identify the most appropriate optimization method to solve (3.2), three well-known methods, i.e., ICM, SA, and GC are compared.

3.3.2 CP Unary Potential via WMM

The unary potential ψ_u in CRF-WMM is calculated by taking into consideration the statistical distribution of the CP SAR data. The unary potential in the CRF model is used

to address the class-dependent information of different classes and is defined as:

$$\psi_u(x_i, y_i) = -ln(p(y_i|x_i)) \tag{3.3}$$

where $p(y_i|x_i)$ is designed based on the statistical characteristics of the data. To achieve this, a GMM is typically used [104,117]. However, because the coherency matrix in (2.5) follows a complex Wishart distribution, a new mixture model based on the Wishart distribution is proposed.

The ith pixel in CP SAR imagery is represented by a complex matrix $\mathbf{A}_i = L\mathbf{J}_i$. Because $\mathbf{J}_i$ follows a complex Wishart distribution, the class-dependent distribution of $\mathbf{A}_i$ given $\mathbf{J}_{x_i}$ has the following expression:

$$p(\mathbf{A}_i|\mathbf{J}_{x_i}) = \frac{|\ \mathbf{A}_i\ |^{L-q}\ exp(-tr(\mathbf{J}_{x_i}^{-1}\mathbf{A}_i))}{K(L,q)\ |\ \mathbf{J}_{x_i}\ |^{L}} \tag{3.4}$$

where $\mathbf{J}_{x_i}$ is the average of the covariance matrices of pixels from class x_i. The symbols $|\ .\ |$ and tr denote the determinant and trace operations, respectively. Also, $K(L,q)$ was defined in (2.10). By substituting (3.4) into (3.3), the unary potential for the complex Wishart distribution is represented as:

$$\psi_u(x_i, y_i) = Lln\ |\ \mathbf{J}_{x_i}\ | + ln(K(L,q)) + tr(\mathbf{J}_{x_i}^{-1}\mathbf{A}_i) - (L-q)ln(|\ \mathbf{A}_i\ |) \tag{3.5}$$

After substituting $\mathbf{A}_i = L\mathbf{J}_i$ into (3.5) and eliminating elements that are not a function of x_i, the unary potential in CRF-WMM will be:

$$\psi_u(x_i, y_i) = ln\ |\ \mathbf{J}_{x_i}\ | + tr(\mathbf{J}_{x_i}^{-1}\mathbf{J}_i) \tag{3.6}$$

3.3.3 CP Pairwise Potential

The pairwise potential ψ_p in (3.1) represents the relationships between the labels and the features of neighbouring pixels. ψ_p in CRF-WMM is defined by a new similarity measure designed based on the properties of CP SAR data. ψ_p has been implemented using the following expression:

$$\psi_p(x_i, x_j, y_i, y_j) = \lambda(y_i, y_j).\mu(x_i, x_j) \tag{3.7}$$

where $\mu(x_i, x_j)$ is considered as the multi-level logistic (MLL) model [118], and $\lambda(y_i, y_j)$ is a measure of feature similarity. Designing a $\lambda(y_i, y_j)$ in ψ_p for CP data is necessary to achieve the best performance since doing so can more effectively constrain pixels with a

strong spatial correlation to have the same class label while also reducing the effects of noise and spatial heterogeneity in SAR imagery leading to more accurate predictions.

The similarity measure $\lambda(y_i, y_j)$ in (3.7) returns a larger value for two pixels with a stronger spatial correlation. Although a metric defined based on the full CP information theoretically can better reveal the similarity among pixels, using all elements in the covariance matrix in (2.6) will increase the computational burden. We choose J^{22} because it is suitable for oil spill detection due to its higher signal-to-noise ratio and sensitivity to the sea surface roughness [36]. Therefore, a similarity measure is designed based on the J^{22} elements given in (2.6):

$$\lambda(J_i^{22}, J_j^{22}) = exp(\frac{- \mid J_i^{22} - J_j^{22} \mid^2}{2\theta^2}) \tag{3.8}$$

where θ controls the scale of the Gaussian kernel. $\lambda(J_i^{22}, J_j^{22})$ generates high values if $\mid J_i^{22} - J_j^{22} \mid^2$ is small; therefore, $\lambda(J_i^{22}, J_j^{22})$ constrains pixels with similar values to each have the same class label. By substituting (3.8) in (3.7), the pairwise potential in CRF-WMM will be:

$$\psi_p(x_i, x_j, y_i, y_j) = exp(\frac{- \mid J_i^{22} - J_j^{22} \mid^2}{2\theta^2}).(1 - \delta(x_i, x_j)) \tag{3.9}$$

3.3.4 Optimization Approaches

Solving the CRF is an ill-posed problem; therefore, an optimization algorithm is necessary to solve it. To assess the variety of optimization algorithms, it is necessary to evaluate their performance in oil spill candidate detection by using CP SAR data. The MAP problem in (3.2) can be reformulated as:

$$\hat{X} = arg\min_X(-ln(p(X|Y))) \tag{3.10}$$

By substituting the unary potential (3.6) and the pairwise potential (3.9) into (3.10), the MAP problem can be expressed as:

$$\hat{X} = arg\min_X\{\sum_{i\in S}(ln \mid \mathbf{J}_{x_i} \mid + tr(\mathbf{J}_{x_i}^{-1}\mathbf{J}_i)) + \beta\sum_{i\in S}\sum_{j\in N_i} exp(\frac{- \mid J_i^{22} - J_j^{22} \mid^2}{2\theta^2}).(1 - \delta(x_i, x_j))\} \tag{3.11}$$

The parameter to be estimated, $\hat{X}$, is a binary variable. So, (3.11) defines a combinatorial optimization problem, which can be solved by several techniques. In this study, we

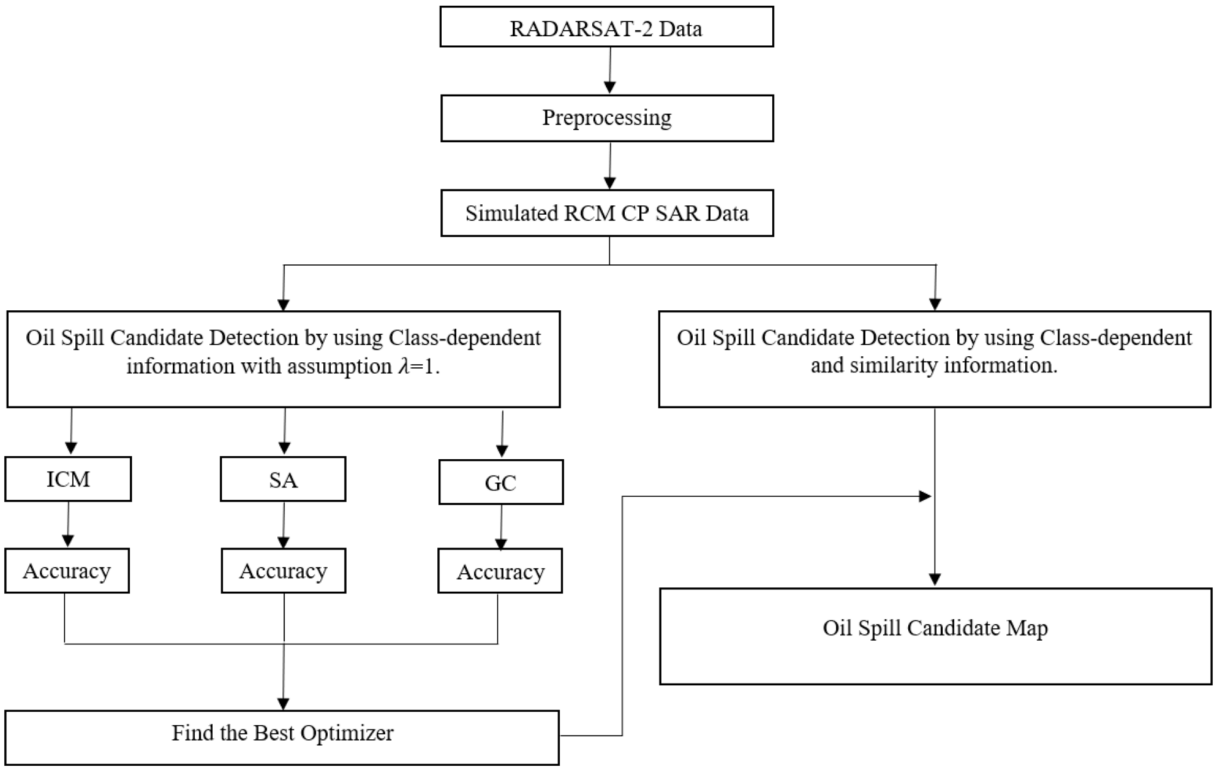


Figure 3.1: The flowchart of the proposed method to detect oil spill candidates from simulated RCM CP SAR data.

investigate and compare ICM, GC, and SA as optimization methods to obtain $\hat{X}$.

3.3.5 Summary of the Proposed Method

Fig. 3.1 illustrates the flowchart of the proposed CRF to detect oil spill candidates. Below we provide a summary of the CRF-WMM model which is a thresholding guided segmentation approach [104]. Given a CP SAR image Y, the following steps are performed:

1. Perform binary thresholding on J^{22} from Y to obtain an initial estimation of X, and use this estimate to calculate the class probabilities for oil-free water $\mathbf{J}_{x_0}$ and oil spills $\mathbf{J}_{x_1}$.

2. Calculate the unary potential as per (3.6).

Table 3.1: The details of the Fine Quad (FQ) single look complex RADARSAT-2 imagery used to simulate CP SAR data. The first column shows the date of acquisition in the format YearMonthDay. The first two images are used to simulate sub-images and the last image is used to simulate a full scene.

Scene	Beam	Acquisition time (UTC)	Range resolution (m)	Azimuth resolution (m)
20100131	FQ6	14:09:13	4.73	4.70
20100224	FQ6	14:09:15	4.73	4.70
20090919	FQ18	02:00:52	4.73	4.95

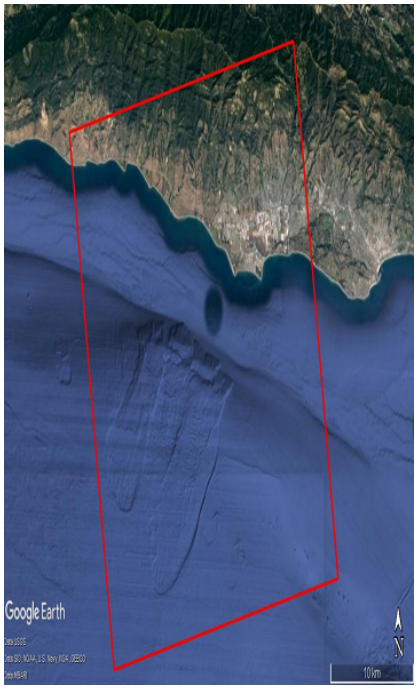
3. Calculate the pairwise potential based on the similarity measure as per (3.9).

4. Calculate the objective function as per (3.11), and solve it using each optimization technique (i.e., ICM, SA, and GC).

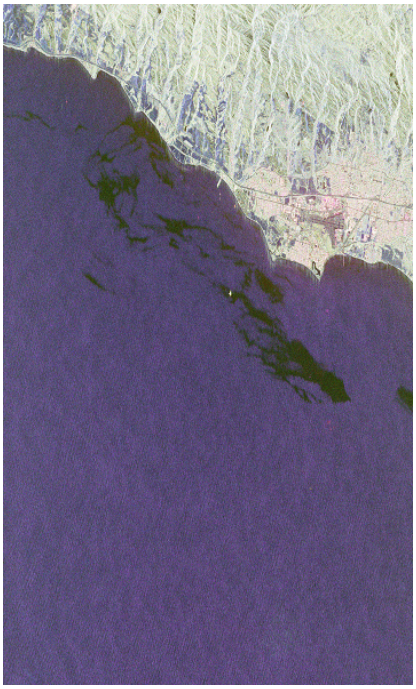
3.4 Study area

In this study, the potential of the proposed CRF in detecting oil spill candidates is investigated by using simulated RCM CP SAR images. To simulate RCM CP SAR images, three QP RADARSAT-2 images acquired over Coal Oil Point, near Santa Barbara, California, USA are used. The details of the RADARSAT-2 images are represented in Table 3.1. These images have been provided by the Canadian Ice Service (CIS) under the Integrated Satellite Tracking of Pollution (ISTOP) program [119]. To simulate CP SAR data, the RCM CP simulator is used. This simulator was developed at the Canada Centre for Mapping and Earth Observation (CCMEO) [8]. Data was simulated for the RCM medium resolution beam mode with 50 m range $\times$ 50 m azimuth nominal resolution and -22 dB noise floor (noise-equivalent sigma zero (NESZ) value). A 9$\times$9 boxcar filter was applied to all the scenes to reduce the effects of speckle noise. A consequence of NRCS values lower than the NESZ is the loss of information. Although utilizing spatial context information and the coherency matrix of CP SAR data along with an appropriate statistical model make the proposed method able to distinguish oil spill candidates from oil-free water with NRCS above the NESZ, characterizing oil spills for NRCS values lower than -22 dB may not be possible [35]. Fig 3.2 shows the location and Pauli color composite of Scene 20090919.

(a)



(b)



(c)

Figure 3.2: (a) Location of Scene 20090919, near Santa Barbara, California, USA. (b) The Pauli decomposition of Scene 20090919 with $|\ S_{HH} - S_{VV}\ |$, $|\ S_{HH} + S_{VV}\ |$, and $2|\ S_{HV}\ |$.

3.5 Experiment and discussion

This section presents the results obtained by the proposed method using the simulated RCM CP SAR sub-images of different sizes and a simulated RCM CP SAR full scene. Moreover, the performance of the optimization algorithms is compared in this section.

3.5.1 Quantitative Measures

In our experiments, ground truth data is generated based on visual inspection. To evaluate the performance of the proposed CRF, we use the following error metrics [104]:

- Commission error (CE) is defined as the ratio of the number of false predicted oil spill candidate pixels to the number of all predicted oil spill candidate pixels. CE is calculated as follows:
$$CE = \frac{A_E - A_T}{A_E} \tag{3.12}$$
where A_E denotes the number of all predicted oil spill candidate pixels and A_T denotes the number of pixels which are correctly classified as oil spill candidates.

- Omission error (OE) is defined as the ratio of the number of false predicted oil-free water pixels to the number of all ground truth oil spill candidate pixels. OE is calculated as follows:
$$OE = \frac{A_R - A_T}{A_R} \tag{3.13}$$
where A_R indicates the total number of ground truth oil spill candidate pixels.

- Averaged error (AE) is the average of CE and OE. AE measures the balanced detection capability of different methods and it is defined as:
$$AE = \frac{CE + OE}{2} \tag{3.14}$$

In this study, we use a grid search on J^{22} to find the best values for β and θ in (3.11). Both parameters are varied from 0.5 to 5 in increments of 0.5, and the result with the lowest AE is chosen as the output of each method.

3.5.2 Experiment with Sub-images

Since oil spills are rare phenomena and usually appear in only small sections of a scene, methods are usually tested using a sub-image containing oil spills [104]. Therefore, to evaluate the performance of the proposed method, we use the two simulated scenes to extract five sub-images of different sizes (see Table 3.1).

Because the CRF-WMM model consists of three building blocks, i.e., the unary potential, the pairwise potential, and the optimization algorithm, we use ablation experiments to isolate their roles and test their importance. First, to justify our unary potential, we compare the proposed WMM with the GMM model built on a vector of J^{11}, $abs(J^{12})$, and J^{22}. That is a common approach [65, 104, 117]. Second, to justify the use of the pairwise potential in the CRF-WMM over the traditional isotropic homogeneous MRF model commonly used [106, 117], we compare the proposed pairwise potential with an MRF potential implemented by discarding the similarity measure $\lambda(y_i, y_j)$ in (3.7) and only use $\mu(x_i, x_j)$. In this experiment, we also compare different optimization methods to identify the one which achieves the highest accuracy.

Unary Potential and Optimization Approaches

Fig 3.3 shows the results of different methods, i.e., the combination of different unary implementations (GMM vs. WMM) and the different optimizers (ICM, SA, and GC). Regardless of the optimization techniques used, the GMM model tends to incorrectly classify oil spill candidate pixels as oil-free water. However, WMM is able to reduce the number of false negatives. This is because the phase information helps the model more appropriately characterize the statistical behavior of RCM CP SAR data. Table 3.2 shows the mean values of CE, OE, and AE of the five sub-images by using the different methods. The WMM-based methods achieve much lower OE than the corresponding GMM-based methods. Low OE is more important than low CE for oil spill candidate detection problems because omitted potential oil spills will never be detected in the further classification of the true oil spills and the look-alikes. These results demonstrate the importance and improvements of using the proposed WMM approach for implementing the unary potential in the CRF model.

Comparing the performance of the optimization algorithms, SA outperforms ICM. We hypothesize that this is because the solution for ICM is known to routinely become trapped in a local minimum [120]. However, SA employs a random search scheme to overcome this problem. Although utilizing GC increases CE, it significantly reduces OE in both WMM

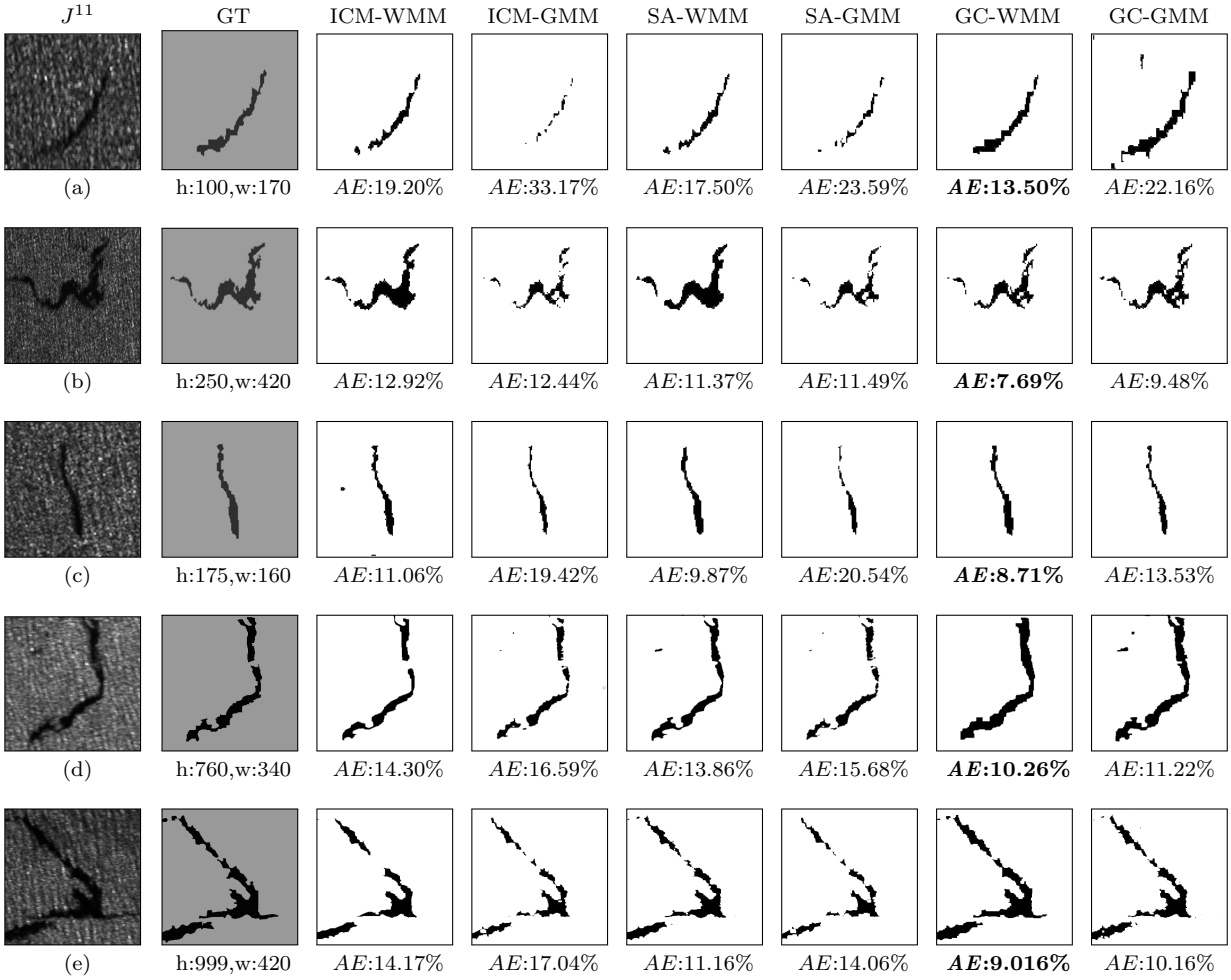


Figure 3.3: Detected oil spill candidates by using the Wishart mixture model (WMM) and Gaussian mixture model (GMM) as the unary potential with different optimization methods and the assumption that $\lambda(J_i^{22}, J_j^{22}) = 1$. The first column shows the five data sets. The second column shows the ground truth (GT) data generated manually based on visual inspection. The h and w stand for the number of pixels in height and width of the sub-images, respectively. Numbers in **bold** indicate the lowest average error (AE) in each sub-image.

and GMM models. We expect this is due to the fact that GC is able to preserve relevant sharp discontinuities while enforcing the piecewise smoothness [121].

Table 3.2: Mean values of the errors for the sub-images achieved by WMM and GMM unary potentials using the different optimization method.

	ICM-WMM	ICM-GMM	SA-WMM	SA-GMM	GC-WMM	GC-GMM
$CE(\%)$	11.46	**1.44**	11.32	2.83	14.04	18.09
$OE(\%)$	17.20	38.02	14.97	31.53	**5.63**	9.40
$AE(\%)$	14.27	19.81	12.75	17.18	**9.83**	13.31

Numbers in **bold** indicate the lowest error using each of the optimization methods

Pairwise Potential

So far, we have considered pairwise potentials, which depend only on the labels of neighboring pixels, while the similarity among features of pixels was ignored. Also, the benefits and improvement of using the proposed WMM approach over GMM for implementing the unary potential in the CRF model was demonstrated, and GC was identified as the optimizer that achieved the highest accuracy. This experiment further demonstrates the role and importance of the similarity measure $\lambda(y_i, y_j)$ defined in (3.8) by incorporating it into the previous GC-GMM and GC-WMM models. Fig 3.4 shows the results obtained by the CRF-WMM model and the baseline CRF-GMM model on the five sub-images using the GC optimizer. The similarity measurement assists the model in preserving the boundaries with higher accuracy by removing wrongly classified pixels. This effect of the similarity measurement is more noticeable for the first, third, and fourth data sets (Fig 3.3 (a, c, d) and Fig 3.4 (a, c, d)). Moreover, employing the feature similarity among neighborhood pixels reduces the number of false positives. For example, some dark areas were classified wrongly as oil spill candidates by GC-GMM in the first dataset but applying $\lambda(y_i, y_j)$ assigns them to the correct class label.

Table 3.3 illustrates the statistics achieved by CRF-WMM and CRF-GMM on the five sub-images. As expected, the performance of the proposed method in all data sets is better than that of CRF-GMM. Compared to the results presented in Table 3.2, there is a significant reduction in the mean values of CE. Specifically, the error is reduced from 14.04% using GC-WMM to 9.04% using CRF-WMM and from 18.09% using GC-GMM to 13.46% using CRF-GMM. This confirms that using $\lambda(y_i, y_j)$ for implementing the pairwise potential in the CRF model reduces the number of false positives. Moreover, CRF-GMM improves OE by about 2%, while CRF-WMM increases OE by about 0.5%. Comparison of the results in Fig 3.3 with those in Fig 3.4 demonstrates consistency with the quantitative measures presented in Table 3.3.

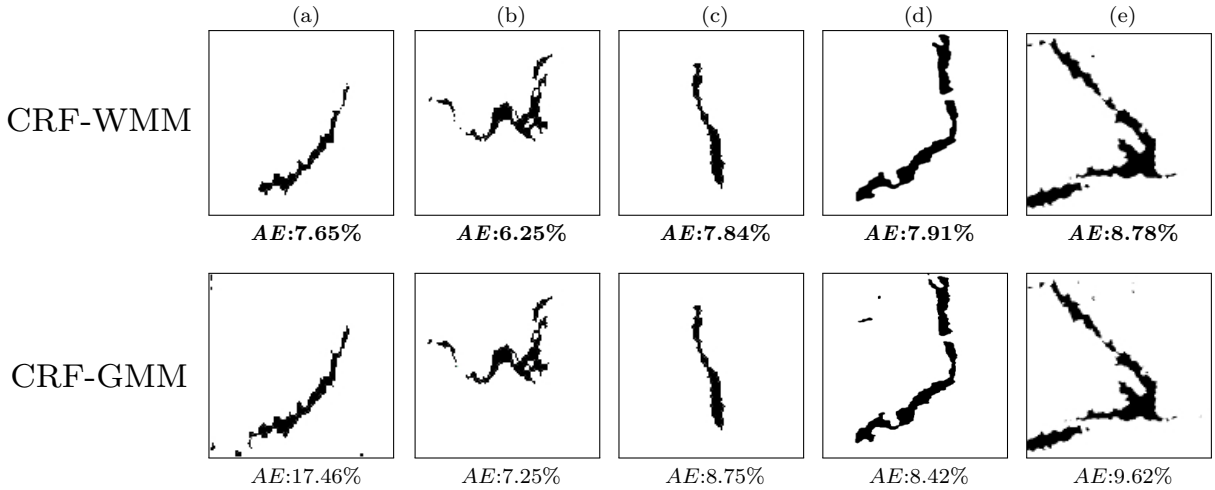


Figure 3.4: Oil spill candidates detected by using the CRF-WMM and CRF-GMM models using the GC optimizer along with their average error (AE). The first and second rows show the results obtained by the CRF-WMM and CRF-GMM models, respectively. Each column shows the sub-images. Numbers in bold indicate the lowest AE in each sub-image

Table 3.3: Mean values of the errors for the sub-images achieved by CRF-WMM and CRF-GMM using the GC optimization method.

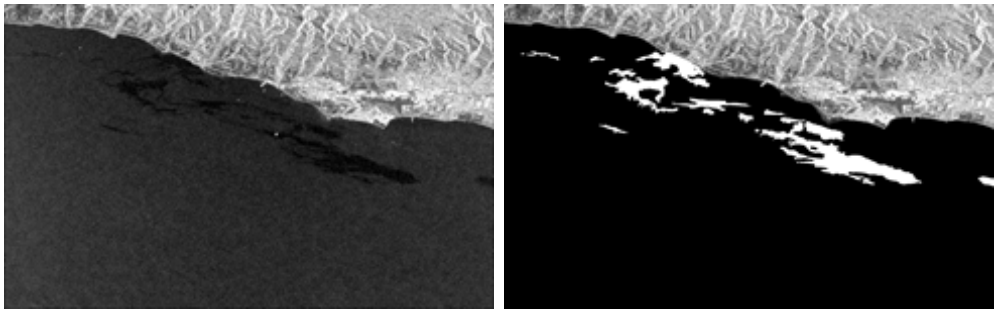
	$CE(\%)$	$OE(\%)$	$AE(\%)$
CRF-WMM	**9.04**	**6.29**	**7.68**
CRF-GMM	13.46	7.13	10.30

Numbers in **bold** indicate the lowest error using each of the optimization methods

3.5.3 Experiment with a Simulated CP SAR Full Scene

To show that the proposed method can reliably detect oil slicks, we examined the performance of the proposed CRF over a full scene. We used Scene 20090919 to simulate an RCM CP SAR scene. Fig 3.5 (a) illustrates the J^{11} image of the study area. It contains three classes, namely, open water, oil spill candidate, and land area. However, since the problem is a two-class classification, a mask is used to exclude the land areas from the calculation. Fig 3.5 (b) shows the manually generated ground truth oil spill candidate data.

Fig 3.6 shows the results obtained by the different methods and optimization algorithms. As expected, the performance of WMM on the simulated scene of CP SAR data is better than that of GMM. The results obtained by the GMM-based methods are less accurate

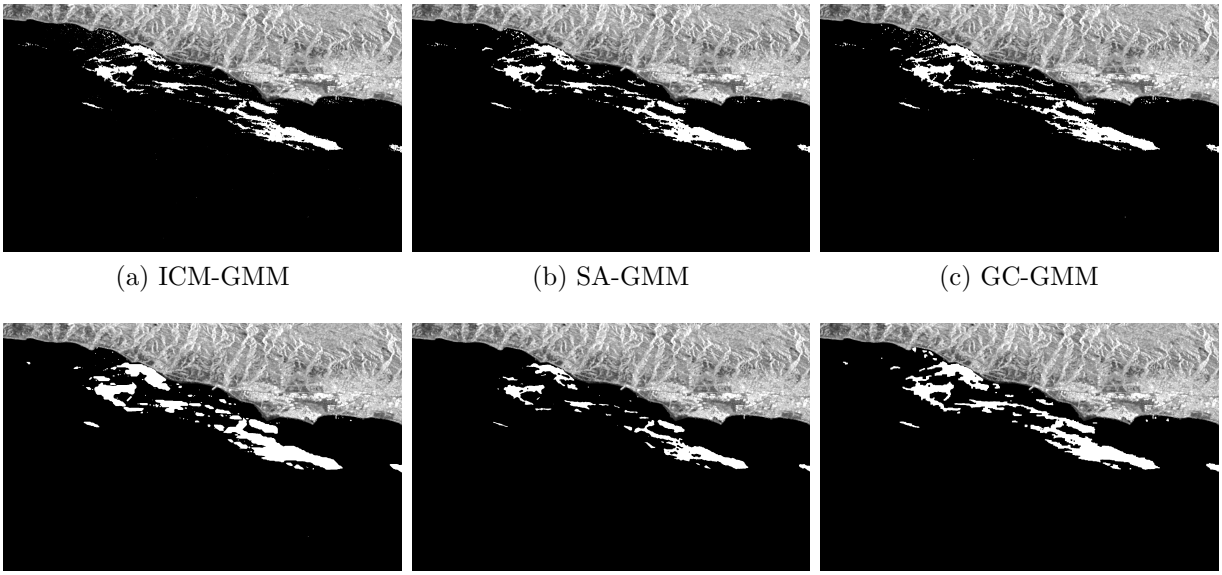


(a) (b)

Figure 3.5: (a) 50 m resolution RADARSAT Constellation Mission right-horizontal polarization imagery (J^{11}) of an area near Santa Barbara simulated from data acquired on 19 September 2009. (b) Ground truth data of oil spill candidates generated manually based on visual inspection.

due to more false positive being detected compared to the WMM-based methods, specially closer to the coastline. Among WMM-based methods, the ICM-WMM method tends to blur the boundaries by accepting pixels close to the boundaries as oil spill candidates. The SA-WMM and GC-WMM methods effectively identify oil-free water even when the backscatter of the open water has high variability. This is because, compared to GMM-based methods, they utilize a more appropriate statistical model. Table 3.4 indicates the errors obtained by the proposed and baseline GMM methods. An examination of Table 3.4 shows that the values of OE obtained by WMM-based methods are much lower than those obtained by GMM-based ones. A high rate of CE causes a high computational cost in removing look-alike candidates. Thus, the capability of a model to produce a balanced OE and CE error is also important. Compared to GMM-based methods, WMM-based methods reach lower AE. This confirms the balanced detection capability of the proposed method. Moreover, as per the results in Subsection 3.5.2, the accuracy of the models using GC is higher than that of the other optimization algorithms. In the next experiment, the performance of CRF-WMM and CRF-GMM on the simulated full scene is evaluated. Since GC achieves a higher accuracy for oil spill detection than the other algorithms, the GC optimization method is used.

Fig 3.7 illustrates the results obtained by CRF-WMM and CRF-GMM. Visually, the similarity measure reduces the number of misclassified pixels in both CRF-WMM and CRF-GMM. Furthermore, CRF-GMM achieves a smaller number of false positives near to the coastline. The statistics of the numerical measures achieved by CRF-GMM, and



(a) ICM-GMM (b) SA-GMM (c) GC-GMM

(d) ICM-WMM (e) SA-WMM (f) GC-WMM

Figure 3.6: Oil spill candidate detection based on the WMM and GMM unary potentials by using the different optimization methods and assumption that $\lambda(J_i^{22} - J_j^{22}) = 1$.

Table 3.4: Values of the errors for the images achieved by WMM and GMM unary potentials using the different optimization method.

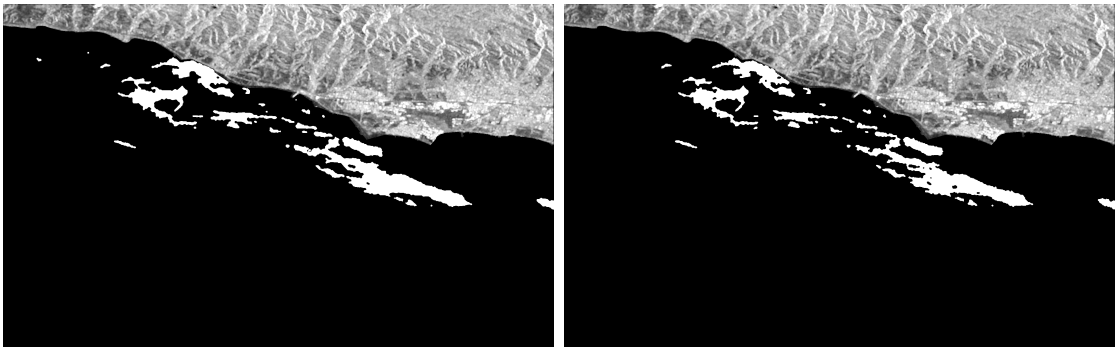
	ICM-WMM	ICM-GMM	SA-WMM	SA-GMM	GC-WMM	GC-GMM
$CE(\%)$	23.15	**4.94**	14.34	6.06	13.85	12.72
$OE(\%)$	9.54	31.87	13.11	24.11	**9.08**	16.44
$AE(\%)$	16.34	18.41	13.73	15.07	**11.58**	14.58

Numbers in **bold** indicate the lowest error using each of the optimization methods

Table 3.5: Values of the errors for the simulated full scene achieved by CRF-WMM and CRF-GMM using the GC optimization method.

	$CE(\%)$	$OE(\%)$	$AE(\%)$
CRF-WMM	**4.51**	**14.61**	**9.56**
CRF-GMM	4.55	18.03	11.29

Numbers in **bold** indicate the lowest error using each of the optimization methods



(a) CRF-WMM (b) CRF-GMM

Figure 3.7: Oil spill candidate detection based on the CRF-WMM and CRF-GMM.

CRF-WMM models are shown in Table 3.5, which are consistent with the visual detection results. The values of average error are 9.56% and 11.29% for CRF-WMM and CRF-GMM, which are almost 2% and 3% lower than those obtained by GC-WMM and GC-GMM, respectively.

3.6 Conclusion

This chapter presented a CRF-WMM method tailored to the CP SAR statistics to utilize both the full CP information and the spatial information in CP SAR imagery for enhanced oil spill candidate detection. First, to utilize the statistical properties of CP SAR data, we designed a unary potential based on the complex Wishart distribution. Second, to take advantage of spatial information, we included a similarity measure based on J^{22}. Theoretically, J^{22} is sensitive to the sea surface roughness and has a high signal-to-noise ratio. The empirical results prove that this similarity measure can improve the accuracy of detecting oil spill candidates. Finally, to solve the proposed ill-posed CRF model, we utilized three common optimization algorithms and compared their performance to specify the most appropriate optimization algorithm for detecting oil spill candidates in simulated RCM CP SAR data. The proposed approach is tested on both the simulated sub-images and the full scene. The results demonstrate that the proposed CRF can better delineate oil spill candidates than the traditional CRF and MRF approaches that do not consider the properties of the CP data. Overall, the proposed CRF can delineate oil spill candidates without being significantly affected by oil-free water and oil spill heterogeneities. In addition, the number of false negatives in the CRF-WMM model is much lower than

that in the other approaches, meaning that using the proposed CRF can decrease the risk of misclassifying oil spill candidate pixels that will not be detected in oil spill classification methods. Given the limitations of hand-crafted features in addressing signature ambiguity, as they may not provide a comprehensive representation of all classes and are susceptible to noise and changing conditions, it is crucial for future research to explore deep learning methods that incorporate feature learning approaches. These feature learning techniques have the potential to overcome the challenges posed by signature ambiguity and enhance the robustness and adaptability of classification algorithms in various conditions.

Chapter 4

Region-Based Sea Ice Mapping using Compact Polarimetric Synthetic Aperture Radar Imagery with Learned Features and Contextual Information

Operational sea ice maps are usually generated manually using DP SAR satellite imagery but there is strong interest in automating this process. Existing sea ice scene classification algorithms using CP imagery rely on hand-crafted features while neural networks offer the potential of features that are more discriminating. We have developed a new and effective sea ice classification algorithm that leverages the nature of CP data. First, a residual-based convolutional neural network (ResCNN) is implemented to classify each pixel. In parallel, an unsupervised segmentation is performed to generate regions based on CP statistical properties. Regions are assigned a single class label by majority voting using the ResCNN output. For testing, QP SAR sea ice scenes from the RADARSAT-2 are used and QP, DP, CP, and reconstructed QP (RQP) modes are compared for classification accuracy, while also comparing to other classification approaches. Using CP achieves an overall accuracy of 96.86%, which is comparable to QP (97.16%), and higher than RQP and DP data by about 2% and 10%, respectively. The implemented algorithm using CP imagery provides an improves option for automated sea ice mapping. This chapter is based on the published paper [57].

4.1 Introduction

Sea ice maps are essential for applications such as climate change interpretation and ocean navigation [12, 122]. SAR from satellites is the primary source of imagery used to generate sea ice maps. National ice centers, such as the Canadian Ice Service (CIS), rely on trained operators to manually generate sea ice maps primarily using SAR imagery, a time-consuming process. Automated sea ice classification methods using SAR imagery have been sought for decades [43].

The RCM CP operates in CTLR (right circular transmit and linear receive) mode [9, 123] and generates swaths (500 km) comparable to DP imagery while preserving the phase information between channels, thereby providing more comprehensive scene information than DP [3, 40]. QP swaths are much smaller (20 km) and are not viable for operational ice mapping. Compared to the more extensive literature for sea ice mapping using DP, there exist limited publications that assess CP SAR data for automated sea ice mapping [4, 39–46]. Previous sea ice classification methods using CP utilize hand-crafted features such as intensity images, polarimetric features, and texture features. Some classification methods use RQP data derived from CP SAR to leverage well-known QP parameters whose usage and interpretation are known [4, 43]. Existing CP SAR sea ice classification methods, although successful, have limitations.

Using hand-crafted features have limitations because these cannot be assured to be a comprehensive representation of all classes and are sensitive to noise and changes in conditions [124]. In addition, feature selection processes are time-consuming and may lead to the loss of important information or strong correlations between features [125]. In contrast, feature learning methods allow a system to automatically extract effective features for specific data and conditions without human intervention [48]. CNNs have been used to learn features from SAR imagery in support of ice concentration estimation [126], ice-water classification [127–131], sea ice change detection [132], and ship-iceberg discrimination [133]. ResCNN methods are able to learn effective feature representations and are highly adaptable to various tasks and datasets [134]. We are not aware of published research that uses deep learning models applied to CP imagery to support sea ice mapping.

The complex Wishart distribution is a known statistical property of the multilook CP coherence matrix, arising from the complex Gaussian distribution of the backscattered field in a CP SAR data [135]. Given this knowledge, it would be advantageous to include the Wishart distribution as part of the CP sea ice classification approach to account for the nature of the data. Furthermore, the use of spatial context enhances both the classification accuracy and the algorithmic efficiency [136]. Therefore, the integration of spatial context should be considered in generating sea ice maps.

We propose a new region-based automated sea ice classification methodology that incorporates learned features, spatial context, and statistical properties of various SAR modes. Uniquely, we apply the deep learning algorithm directly to the CP data to generate a pixel-level classification. As a novel contribution, CP will be compared to DP, RQP, and QP with regards to generating accurate sea ice maps. Identifying the most effective mode for SAR sea ice classification supports improved operational algorithmic capability.

An ResCNN model [134] is used to obtain pixel-level sea ice maps by learning sea ice features from CP SAR data. An existing unsupervised semantic segmentation based on statistical characteristics of the CP data [137] is applied to obtain homogeneous and edge-preserved regions. Similar approaches, using appropriate statistical distributions, are implemented for each of the SAR modes under consideration. To generate region-based sea ice maps, as supported by previous studies [43, 49], a majority voting process is employed to combine pixel-level classified and segmented images.

Experiments are based on a pair of RADARSAT-2 QP scenes that are used to simulate corresponding RCM CP scenes. The region-based sea ice classification approach achieves accurate sea ice maps and demonstrates that CP can achieve comparable sea ice classification performance relative to QP and outperform DP and RQP.

A literature review of the sea ice classification methods using CP SAR data is presented in Section 4.2. The proposed method is explained in Section 4.3. The experiments and the corresponding results are shown and analyzed in Section 4.4, and the conclusions are presented in Section 4.5.

4.2 Background

In general, classification algorithms utilizing CP can be divided into two categories: algorithms using features derived directly from CP data and algorithms employing RQP data derived from CP data. Each of these categories will be discussed next.

In the first category, polarimetric features such as $m - \chi$ decomposition features [116] are extracted from the Stokes vector of CP data. Perhaps the first evaluation of CP imagery's ability to differentiate open water from sea ice was visually conducted by Charbonneau *et al.* [40]. They used simulated CP imagery data from QP aircraft-based SAR images acquired over the Canadian Arctic. Subsequently, several studies then evaluated the capability of CP polarimetric features extracted from the Stokes vector in distinguishing different ice types and open water [39–42, 45, 138].

In the second category, rather than using CP data directly, the QP covariance matrix is reconstructed from CP data [74, 75] and used with QP scene classification methods. Zhang *et al.* [44] evaluated the ability of CP modes to reconstruct QP information and they recommended using CTLR mode for studying sea ice classification. Ainsworth *et al.* [139] demonstrated that RQP's potential in classifying crop fields is comparable to that of CP.

In addition to using backscatter measures, hand-crafted features can be used to augment the feature set for a sea ice classification task. However, using hand-crafted features requires domain expertise and selecting the proper parameters to generate effective features is time-consuming [140]. Deep learning methods, in contrast, are data-driven and do not require prior knowledge or assumptions, and can automatically learn features from data [140, 141] resulting in removing the need for manual feature engineering [141]. Thus, it is highly advantageous to evaluate the potential of CP imagery for generating sea ice maps using deep learning techniques.

Combining sea ice maps generated by pixel-level classification with segmented images improves classification accuracy [43, 49]. Leigh *et al.* [49] utilized IRGS [65] and a glocal approach on RADARSAT-2 images to identify homogeneous regions and an SVM classifier to generate pixel-level sea ice maps using gray-level co-occurrence matrix (GLCM) texture and intensity features. The basic IRGS algorithm [65, 142, 143] uses intensity images with Gaussian statistics in a spatial context model. CP enhances the measurement potential of radar illumination by providing the 2×2 coherence matrix of the backscattered field which follows the complex Wishart distribution [135, 137, 144]. As a result, to fully utilize the information provided by CP data, CP-IRGS [137] should be applied to segment it. Ghanbari *et al.* [43] employed Polarimetric IRGS (PolarIRGS) [145] to segment RQP images with a support vector machine (SVM) pixel-level classification [146] to label regions using CP polarimetric features.

Taking into account these motivations, we propose a region-based ResCNN method that utilizes learned features, spatial context, and statistical characteristics of CP imagery to classify SAR sea ice scenes. The study also evaluates and compares the potential of CP imagery for generating sea ice maps against comparable approaches for DP, RQP, and QP modes. This comparison is necessary as different SAR data modes have distinct characteristics that can impact their ability to generate sea ice maps. By considering these differences, the most appropriate data source can be assessed.

The methods proposed by Ghanbari *et al.* [43] and Leigh *et al.* [49] are used as baselines to compare the performance of the proposed method.

4.3 Methodology

4.3.1 Overview

The proposed region-based classification method consists of the following components shown in Fig. 4.1.

- The various SAR modes (Subsection 4.3.2) each used as source imagery are CP, DP, RQP, and QP as input to the region-based segmentation and pixel-level classification methods.
- A ResCNN model is used to generate the pixel-level sea ice maps by using learned features from each of the modes (Subsection 4.3.3).
- Unsupervised segmentation algorithms, that are mode dependent, are used to generate homogeneous, contiguous regions with accurate class boundaries (Subsection 4.3.4).
- The segmented and pixel-level classified images are combined by a region-based majority voting approach (Subsection 4.3.5).

To ensure consistency and enable comparisons with the baseline methods, we employed the polarimetric features listed in Table 4.1 in the same manner as Ghanbari *et al.*[43]. Meanwhile, the same features as Leigh *et al.*[49] for the RH and RV CP scenes were used. These features are intensity images as well as local average, maximum intensity, and GLCM including applied second moment, contrast, correlation, dissimilarity, entropy, homogeneity, inverse moment, mean, and standard deviation extracted using three window sizes of 3×3, 9×9, and 17×17. The texture features were then averaged across the four dominant directions, known as isotropic GLCM step directions.

4.3.2 SAR Data

The mathematical basis of the four SAR modes (QP, CP, RQP, and DP) used in this study are described here.

Table 4.1: List of CP polarimetric and amplitude features.

Name	Description	# [a]
RH,RV	intensity values of RH and RV channels [147]	2
α	scattering mechanism parameter [148]	1
μ_c	circular polarization ratio [8]	1
u	conformity coefficient [40]	1
ρ	correlation coefficient of RH and RV [40]	1
m	degree of polarization [116]	1
H_i	Shannon entropy, intensity component [40]	1
H_p	Shannon entropy, polarimetric component [40]	1
$m-\chi$	m-chi decomposition of CP data [116]	3
$S_0, ..., S_3$	Stokes vector components [149]	4

[a] # shows the number of features

QP SAR Data

Each pixel in the observed QP imagery is represented by four elements in a 2×2 scattering matrix [73, 150]:

$$\mathbf{S} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \tag{4.1}$$

where S_{ij} is complex and ij indicate the transmitted and received polarizations [73] which can be either horizontal (H) or vertical (V). Then, a 3×3 covariance matrix is calculated assuming reciprocity of monostatic radar [73]:

$$\mathbf{C}_{QP} = \begin{bmatrix} \langle |S_{HH}|^2 \rangle & \sqrt{2}\langle S_{HH}S_{HV}^* \rangle & \langle S_{HH}S_{VV}^* \rangle \\ \sqrt{2}\langle S_{HV}S_{HH}^* \rangle & 2\langle |S_{HV}|^2 \rangle & \sqrt{2}\langle S_{HV}S_{VV}^* \rangle \\ \langle S_{VV}S_{HH}^* \rangle & \sqrt{2}\langle S_{VV}S_{HV}^* \rangle & \langle |S_{VV}|^2 \rangle \end{bmatrix} \tag{4.2}$$

where $\langle ... \rangle$ and * indicate spatial ensemble averaging and the conjugate transpose, respectively. $\mathbf{C}_{QP}$ follows a complex Wishart distribution [73].

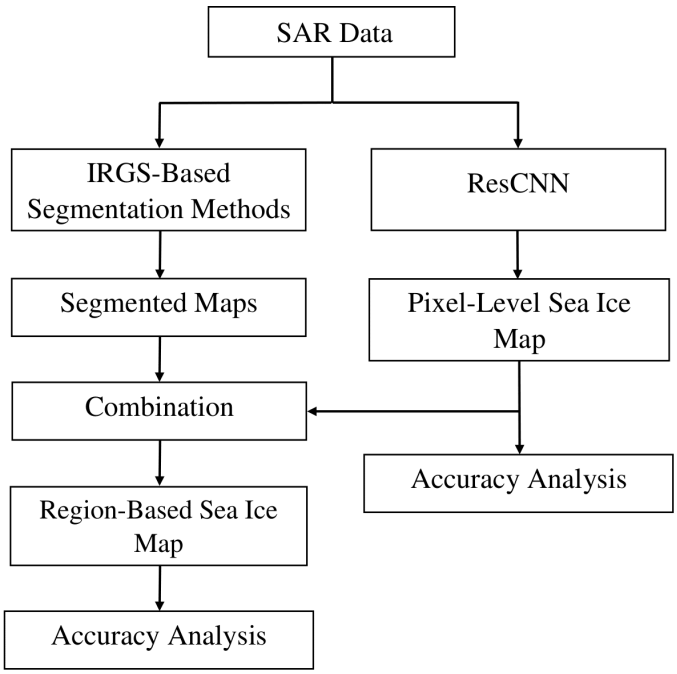


Figure 4.1: Flowchart of the main steps of the proposed classification method.

Simulated CP SAR Data

The coherency matrix of CP SAR data is a 2×2 semi-positive definite Hermitian matrix. For a CTLR mode, the coherency matrix is given as [4]:

$$\mathbf{J}_{CP} = \begin{bmatrix} \langle \mid S_{RH}^2 \mid \rangle & \langle S_{RH} S_{RV}^* \rangle \\ \langle S_{RV} S_{RH}^* \rangle & \langle \mid S_{RV}^2 \mid \rangle \end{bmatrix} \tag{4.3}$$

where R stands for transmitted right circular polarized wave. $\mathbf{J}_{CP}$ follows a complex Wishart distribution [151]. Based on the observed QP data, CP data can be derived. To derive S_{RH} and S_{RV}, the equations $S_{RH} = (S_{HH} - iS_{HV})/\sqrt{2}$ and $S_{RV} = (S_{HV} - iS_{VV})/\sqrt{2}$ are used [9].

Reconstructed QP SAR Data

Based on the CP data, the QP data can be reconstructed to be used as an alternative representation of the CP data. To reconstruct the 3×3 covariance matrix of QP, reflection

symmetry resulting in $\langle S_{HH}S_{HV}^{*}\rangle = \langle S_{HV}S_{VV}^{*}\rangle = 0$ and cross-pol ratio must be used [74]. Therefore, the reconstructed covariance matrix is expressed as [76]:

$$\mathbf{C}_{RQP} = \begin{bmatrix} \langle |S_{HH}|^2 \rangle & 0 & \langle S_{HH}S_{VV}^{*} \rangle \\ 0 & 2\langle |S_{HV}|^2 \rangle & 0 \\ \langle S_{VV}S_{HH}^{*} \rangle & 0 & \langle |S_{VV}|^2 \rangle \end{bmatrix} \tag{4.4}$$

In this paper, to calculate $\mathbf{C}_{RQP}$ elements, an iterative method proposed by Nord *et al.* [75] is used.

DP SAR Data

In contrast to QP and CP, DP data does not include phase information between channels. Therefore, the covariance matrix cannot be calculated. In this study, simulated RCM HH and HV intensity channels are used.

4.3.3 Sea Ice Classification using ResCNN Model

Due to the limited availability of annotated SAR data for sea ice classification tasks, overfitting can occur using deep learning models such as ResNet [134] or VGG [152]. A simpler deep learning model is used to alleviate this problem. In this study, a four-block ResCNN model is used and trained by minimizing the multi-class cross-entropy lost function [153]. As indicated in Table 4.2, each block consists of two convolution operators, and the number of feature maps in each block is set to 32, 48, 64, and 80, respectively. To allow nonlinear expression ability to the ResCNN model, a standard ReLU (rectified linear activation unit) activation function is used [154]. After applying a global average operator on the output of the last block, a fully connected layer is used to map the features into one of K classes.

To minimize the loss function, the Adam optimizer [155] is employed. The distribution of layers is affected by changing the parameters of previous layers in a CNN model [156]. To overcome this limitation, the batch normalization method is used. To help reduce the risk of overfitting, dropout is applied. After training the ResCNN model, all pixels in a SAR image are classified to obtain pixel-level sea ice maps.

4.3.4 Obtaining Homogeneous Edge-Preserved Regions

The effective preservation of boundaries between different ice types when generating sea ice maps is achieved through the use of unsupervised region-based segmentation. The resulting

Table 4.2: Structure of the ResCNN model along with the operators.

Layer name	Output Size	Operators
Block 1	$17 \times 17 \times 32$	$\begin{bmatrix} 3 \times 3 \times 32 \\ 3 \times 3 \times 32 \end{bmatrix}$
Block 2	$9 \times 9 \times 48$	$\begin{bmatrix} 3 \times 3 \times 48 \\ 3 \times 3 \times 48 \end{bmatrix}$
Block 3	$5 \times 5 \times 64$	$\begin{bmatrix} 3 \times 3 \times 64 \\ 3 \times 3 \times 64 \end{bmatrix}$
Block 4	$3 \times 3 \times 80$	$\begin{bmatrix} 3 \times 3 \times 80 \\ 3 \times 3 \times 80 \end{bmatrix}$
Global Average	$1 \times 1 \times 64$	3×3 average pool
Classification	K	$64 \times K$ fully connected
Softmax	K	

regions must meet two criteria: (i) they should be homogeneous, containing only a single class and (ii) they should preserve the boundaries between different classes accurately. The successful IRGS algorithm is chosen because of its effectiveness in segmenting SAR imagery and generating accurate class boundaries [43, 49, 142, 157].

The original IRGS method [65], designed only for SAR amplitude images, was used for the segmentation of sea ice scenes using DP scenes. Yu *et al.* [145] extended IRGS to PolarIRGS and applied it to a land cover-type data set. PolarIRGS leverages all available information in QP by designing a feature model based on the complex Wishart distribution and adapting the spatial context model to better capture the specific characteristics of QP. The edge penalty term is measured using the amplitude images of HH, HV, and VV. The PolarIRGS method is performed here to segment sea ice scenes using $\mathbf{C}_{QP}$ and $\mathbf{C}_{RQP}$.

Ghanbari *et al.* [137] introduced CP-IRGS by similarly modifying IRGS to accommodate CP data. CP scenes are segmented using CP-IRGS and the RH and RV images are used to measure the required edge-penalty term.

4.3.5 Combining Pixel-Based Classification and Region-Based Segmentation

The pixel-level sea ice maps generated by ResCNN are generally accurate because ResCNN learns discriminative features during the training process. However, pixel-level sea ice maps have many errors due to the speckle noise causing individual pixels to be assigned to incorrect classes. Combining the pixel-level classified image with edge-preserving regions results in high-precision sea ice maps. To do that, a majority voting process is applied to each region to determine the sea ice class labels of regions as follows [43]:

$$l_{r_s} = \max_{j=1}^{K} \sum_{i \in r_s} Vote(l_i)_j \tag{4.5}$$

where l_{r_s} indicates the class label of region r_s. K is the total number of ice classes and $i \in r_s$ represents the pixels that comprise r_s. The term l_i indicates label of i^{th} pixel and $Vote(l_i)_j$ is defined as:

$$Vote(l_i)_j = \begin{cases} 1, & \text{if } l_i \text{ belongs to class j} \\ 0, & \text{otherwise} \end{cases} \tag{4.6}$$

4.4 Experiment

This section presents the experiments of the ice-type classification and discusses the performance of the classification in two general cases. In case 1, the potential of CP data in generating sea ice maps using the proposed method is compared against that of DP, RQP, and QP modes. In case 2, the proposed method's performance is compared to that of other baseline methods.

4.4.1 Study Area and Dataset

In this study, two QP RADARSAT-2 scenes are used to generate the RCM DP and CP SAR data and the corresponding RQP SAR data are generated as described in Section 4.3.2. The two QP SAR scenes were acquired in the FQ11 imaging mode with two seconds time difference identified as Scene 56 and Scene 58 over Barrow Strait, located near Somerset Island in the Canadian Arctic collected on May 5, 2010. The range $\times$ azimuth resolutions of the scenes are 5.2×7.6 m. The incidence angle range for the scenes is between 30.20°–32.00°.

Fig. 4.2 (a) illustrates the location of the two scenes. The study area covers approximately 23km by 14km with five different classes: young ice (YI), first-year ice (FYI), multi-year ice (MYI), new ice (NI), and open water (OW) identified by experts in the Canadian Ice Service (CIS). Since the backscatter signatures of OW and NI classes are very similar in Scene 58, they are assumed as the same class (OW/NI) [43]. Fig. 4.2 (b) and (c) show the first element of $\mathbf{J}_{CP}$ of the two simulated CP scenes along with the overlaid labeled pixels.

To simulate CP SAR data from RADARSAT-2 QP scenes, an RCM simulator developed at the Canada Centre for Mapping and Earth Observation (CCMEO) is used [8]. The scenes are simulated for the 30 meters RCM medium resolution beam mode with - 24 dB noise floor (noise-equivalent sigma zero value). A 9×9 boxcar averaging filter is applied to the SAR data to reduce speckle noise. The RCM HH and HV intensity images are also extracted with the same RCM beam mode and averaging filter size.

4.4.2 Training and Testing Data

The number of labeled pixels specified by CIS experts is approximately 1000, which were used to guide the collection of the remaining labeled pixels. CIS experts used the MAGIC software to partition Scene 56 and 58 into a number of grids and assign a label to the central pixel of each grid [158]. Scene 56 is used for training, while Scene 58 is used for testing. Since Scene 56 does not include sufficient numbers of OW/NI samples, 2000 OW/NI samples were obtained from Scene 58 to train models. Table 4.3 shows the number of training and test pixels in each class. The training samples were used to standardize Scenes 56 and 58. $F\times$ 17×17 patches were extracted around each labeled pixel to train the ResCNN models. F stands for the number of input feature maps. In case 1, it corresponds to the absolute values of covariance matrix elements. Thus, F is 2, 3, 4, and 6 for DP, simulated CP, RQP, and QP SAR data, respectively. In case 2, F is 16 and 35 for polarimetric and GLCM feature-based baselines, respectively.

Table 4.3: The number of training and testing pixels for each class.

Name	Description	# of train	# of test
OW/NI	open water and new ice	2000	3367
YI	young ice	5889	6383
FYI	first-year ice	6395	6383
MYI	multi-year ice	5750	5637

The number of training patches in each class is not equal, which is known as an imbalanced dataset problem. To overcome this problem, the data augmentation technique, including horizontal and vertical flips as well as random rotation, is used to expand the number of training patches to 7000 per class.

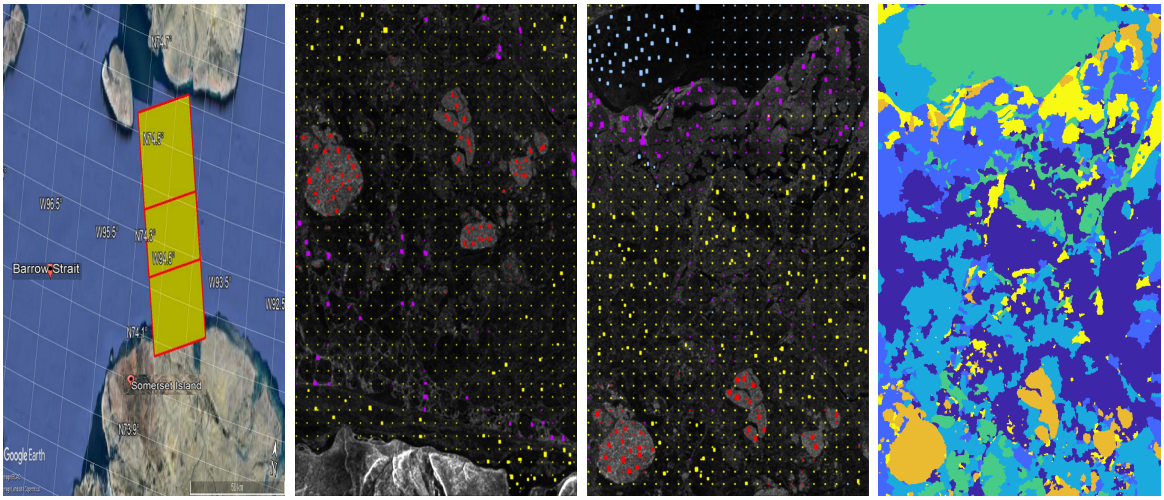
4.4.3 Models Settings

The learning rate, weight decay, and beta parameters in ADAM optimization [155] are set to 1e-5, 0.05, 0.9, and 0.999 in the training phase. The batch size and training epochs are 50 and 200, respectively.

SVM is a machine learning algorithm used previously in support of sea ice mapping [43,49,157,159]. The hyperparameters C, Gamma, and Degree for SVM were explored within the ranges $[10^{-2}, 10^{2}]$, utilizing increments of 10 for the kernel functions of linear, polynomial, and radial basis function. The values of hyperparameters associated with the highest overall classification accuracy (OA) were chosen.

4.4.4 Comparing CP, DP, RQP, and QP Modes

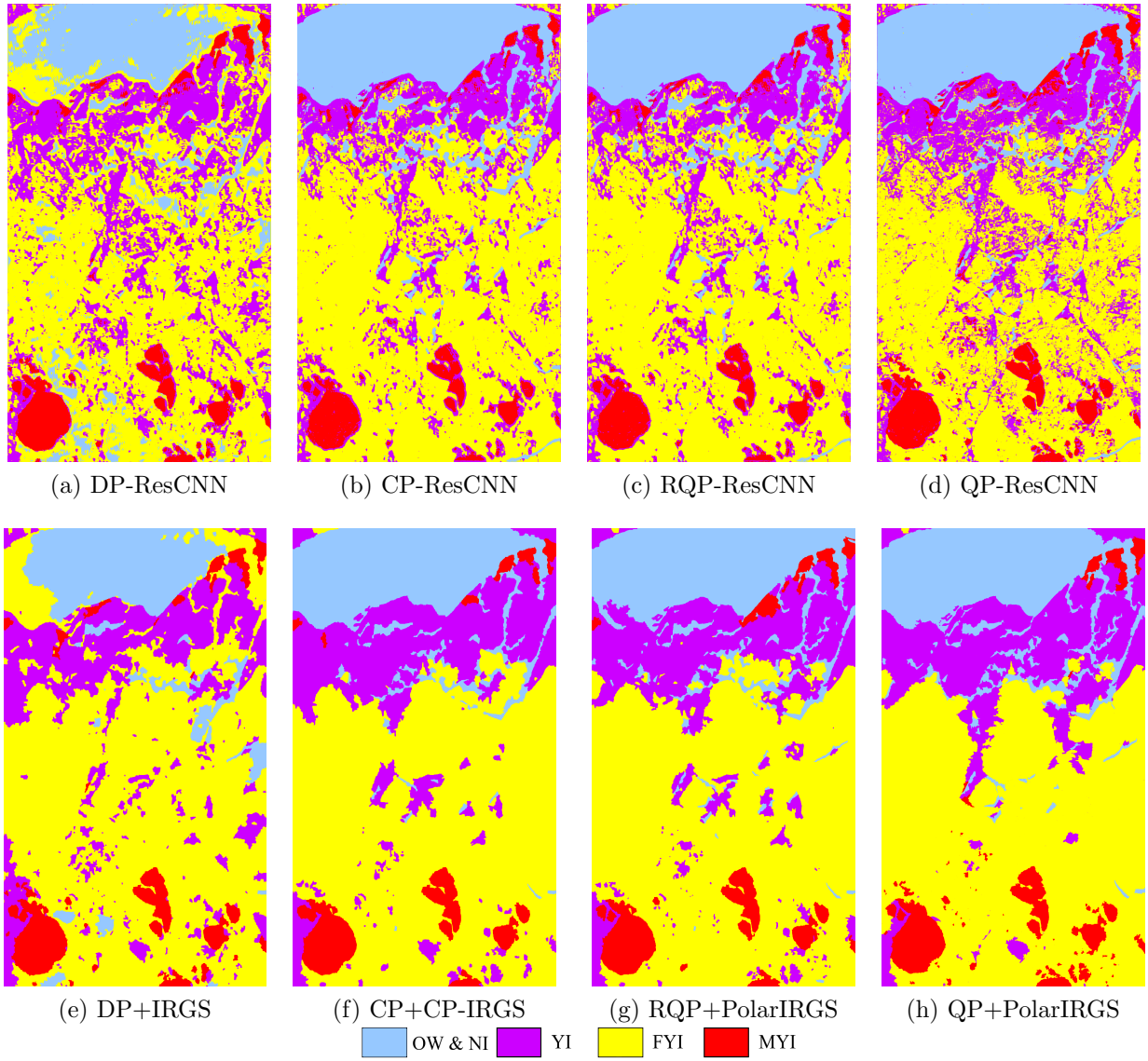
Fig. 4.3 (a)-(d) show pixel-level sea ice maps obtained only by the ResCNN models. The images are resized to fit the page, which may result in some details being lost. In general, the pixel-level sea ice maps appear with many erroneous pixels, probably caused by speckle noise. It is worth noting that all models have mistakenly classified numerous YI pixels in the upper section of the scene as MYI. This misclassification can be attributed to the similarity in intensity values between those YI pixels and the MYI pixels, leading to an inaccurate identification.



(a) Location (b) Scene56 (c) Scene58 (d) IRGS (DP)

(e) Glocal (RV) (f) CP-IRGS ($\mathbf{J}_{CP}$) (g) PolarIRGS ($\mathbf{C}_{RQP}$) (h) PolarIRGS ($\mathbf{C}_{QP}$)

Figure 4.2: (a) The locations of the two RADARSAT-2 QP scenes acquired in the FQ11 imaging mode over Barrow Strait on May 5, 2010. The first element of the covariance matrix of Scenes (b) 56 and (c) 58 along with the overlaid labeled pixels of open water/new ice class (blue), young ice (violet), first-year ice (yellow), and multi-year ice (red). (d)-(h) are the segmentation images obtained using IRGS-based methods using different SAR data mode.



(a) DP-ResCNN (b) CP-ResCNN (c) RQP-ResCNN (d) QP-ResCNN

(e) DP+IRGS (f) CP+CP-IRGS (g) RQP+PolarIRGS (h) QP+PolarIRGS

Figure 4.3: (a)-(d) Pixel-level sea ice maps generated by the ResCNN feature learning classifier using DP (DP-ResCNN), simulated CP (CP-ResCNN), RQP (RQP-ResCNN), and QP SAR data (QP-ResCNN). (e)-(h) segmentation combined with ResCNN classification results. It should be noted that the QP results have been rescaled to match CP data size for presentation purposes.

Table 4.4: Confusion matrices obtained by the region-based ResCNN models using the amplitude scenes of DP, simulated CP, RQP, and QP.

Method		OW/NI	YI	FYI	MYI	User's Accuracy(%)
DP+IRGS	OW/NI	**2299**	6	234	0	90.55
	YI	1	**5053**	366	4	93.16
	FYI	1067	164	**5778**	5	82.38
	MYI	0	1160	5	**5628**	82.85
	Overall Accuracy (%):					**86.16**
	Kappa Coefficient:					**0.8114**
CP+CP-IRGS	OW/NI	**3327**	32	1	0	99.19
	YI	21	**5874**	119	5	97.60
	FYI	19	23	**6261**	7	99.22
	MYI	0	454	2	**5625**	92.50
	Overall Accuracy (%):					**96.86**
	Kappa Coefficient:					**0.9575**
RQP+PolarIRGS	OW/NI	**3307**	8	4	0	99.064
	YI	41	**5655**	287	6	94.42
	FYI	19	27	**6091**	6	99.15
	MYI	0	693	1	**5625**	89.02
	Overall Accuracy (%):					**94.98**
	Kappa Coefficient:					**0.9320**
QP+PolarIRGS	OW/NI	**3342**	14	4	0	99.46
	YI	23	**6155**	324	2	94.63
	FYI	2	26	**6050**	31	99.03
	MYI	0	188	5	**5604**	96.67
	Overall Accuracy (%):					**97.16**
	Kappa Coefficient:					**0.9614**

The QP data generated the highest OA of 89.52%, while CP yielded a slightly lower OA of 88.23% demonstrating that CP data has a comparable ability to provide sea ice maps to QP data. Meanwhile, using RQP achieved an accuracy of 87.24%, which was comparable to CP but also slightly lower than QP. This indicates that RQP did not offer richer information for classification than CP. The lowest OA was obtained by the DP images (80.13%) demonstrating that the polarimetric information in the other three modes

Table 4.5: Confusion matrices obtained by the region-based baseline methods.

Method		OW/NI	YI	FYI	MYI	User's Accuracy(%)
Ghanbari *et al.* [43]	OW/NI	**3324**	7	1	0	99.76
	YI	41	**4666**	290	3	93.32
	FYI	2	26	**6089**	5	99.46
	MYI	0	1684	3	**5629**	76.94
	Overall Accuracy (%):					**90.52**
	Kappa Coefficient:					**0.8719**
Leigh *et al.* [49]	OW/NI	**3206**	15	4	0	99.41
	YI	2	**5791**	554	7	91.14
	FYI	159	576	**5821**	6	88.71
	MYI	0	1	4	**5625**	99.91
	Overall Accuracy (%):					**93.90**
	Kappa Coefficient:					**0.9171**

are useful in support of accurate classification. As shown in Fig. 4.3 (a), DP-ResCNN misclassified several OW/NI and YI pixels as FYI.

Fig. 4.3 (e)-(h) shows the region-level sea ice maps after combining segmentation and classification maps using majority voting. Overall, the region-level results exhibit well-defined homogeneous regions and less noise classifications when compared to the sea ice maps at the pixel level.

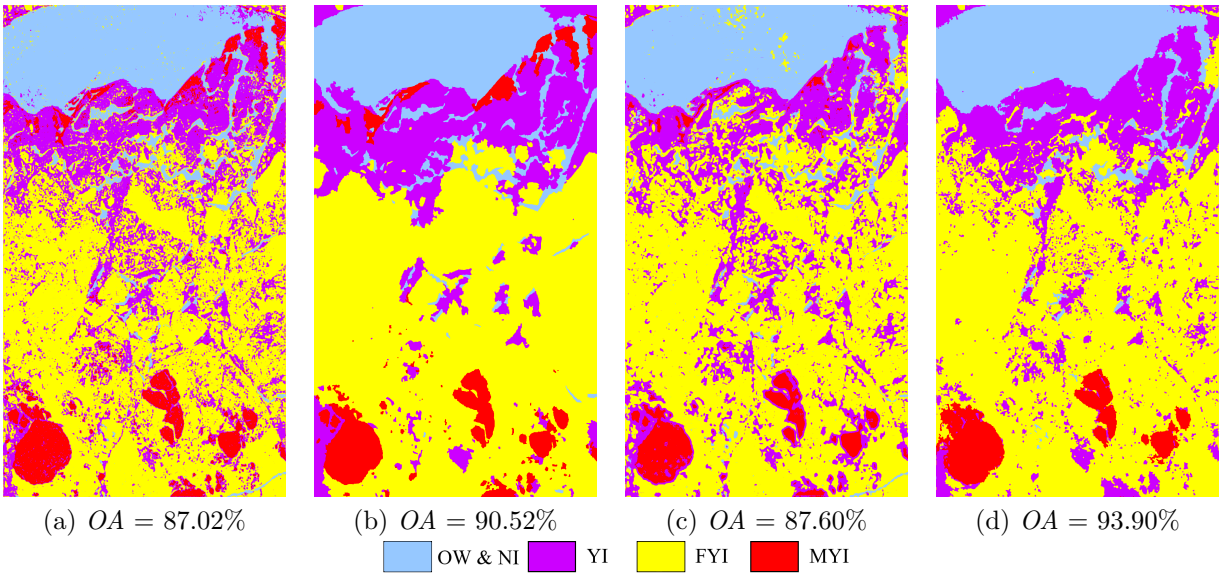
Table 4.4 presents the performance of ice-type classification on the test sampled data. It shows that the combination of QP-ResCNN and PolarIRGS (QP+PolarIRGS) generated the highest OA (97.16%), which was slightly greater than OA of CP-ResCNN combined with CP-IRGS (CP+CP-IRGS) that achieved an accuracy of 96.86%. This demonstrates that CP data can effectively approximate the QP capability for sea ice mapping. The combination of RQP-ResCNN and PolarIRGS (RQP+PolarIRGS) produced an OA of 94.98%, which was lower than CP+CP-IRGS, which does not motivate its use.

The RQP+PolarIRGS method achieved an MYI user's accuracy of 89.02% by misidentifying 10.86% of the YI test samples as MYI, which mainly includes samples in the upper portion of the scene. Both the QP+PolarIRGS and RQP+PolarIRGS methods misclassify several FYI test samples as YI, leading to a decrease of approximately 6% in the YI user's accuracy. In contrast, the CP+CP-IRGS approach can more accurately differentiate between YI and FYI classes.

4.4.5 Performance Comparison of the Proposed Methodology with the Baselines

Fig. 4.4 shows the sea ice maps generated by the baseline methods by Ghanbari *et al.* [43] and Leigh *et al.* [49]. As can be seen in Fig. 4.4 (a) and (c), the baseline method by Ghanbari *et al.* yields many YI samples misclassified as MYI, whereas the number of misclassified YI samples using the approach by Leigh *et al.* is lower. This indicates that using GLCM features enables SVM to distinguish MYI samples from YI more effectively than polarimetric features.

The corresponding confusion matrices are displayed in Table 4.5, which show that the approach proposed by Leigh *et al.* achieving an OA of 93.90% performed better than the method suggested by Ghanbari *et al.*, which obtained an OA of 90.52%. In accordance with Fig. 4.4 (b), the user accuracy of MYI obtained by Ghanbari *et al.* is 76.94% which is indicative of the misclassified YI pixels at the upper part of the scene. While the user accuracy of FYI obtained by Leigh *et al.* is 88.71%, mainly due to the misclassification of



(a) $OA = 87.02\%$ (b) $OA = 90.52\%$ (c) $OA = 87.60\%$ (d) $OA = 93.90\%$

Figure 4.4: Sea ice maps indicating OA for baseline approaches. (a) is pixel-based and (b) is region-based using the method by Ghanbari *et al.* [43] while (c) is pixel-based and (d) is region-based using the method by Leigh *et al.* [49].

YI samples as FYI at the upper part of the scene (Fig. 4.4 (d)).

The approach using ResCNN achieved an 88.23% OA, which was about 1.5% higher than the OA attained by each baseline method. Notably, the ResCNN classifier utilized only amplitude CP scenes, whereas the baseline methods incorporated GLCM and polarimetric features as well. Based on the results presented in Tables 4.4 and 4.5, the region-based sea ice classification approach proposed in this study achieved the highest OA of 96.86% compared to the performance of the baseline methods.

4.5 Conclusion

A ResCNN region-based automated sea ice classification algorithm which utilizes CP SAR data was introduced. The proposed approach incorporates learned features and spatial information and leverages the statistical characteristics of the CP coherence matrix to produce accurate sea ice maps. The experimental results reveal that the proposed method yields sea ice maps with higher accuracy compared to DP and RQP, and performs comparably to QP. These findings suggest that CP data has greater potential than DP for generating sea ice maps, and there is no need to reconstruct QP data from CP. Additionally, the study showed that the potential of CP data in generating sea ice maps is comparable to that of QP.

The ResCNN classifier proposed in this chapter achieved a higher overall classification accuracy (88.23%) compared to the baseline methods, without the need for polarimetric and GLCM features. These results confirm that feature learning classifiers can improve the accuracy of sea ice maps over traditional machine learning methods such as SVM. By achieving the highest overall accuracy of 96.86%, the proposed region-based sea ice classification method outperformed the baseline methods. This highlights the significance of incorporating the statistical properties of CP data and learned features in the sea ice classification process.

While CNNs excel at extracting local features, they inherently lack the ability to capture long-distance dependencies among pixels. This limitation becomes critical for tasks that involve a high degree of spatial heterogeneity in targets, such as land cover classifications. Hence, the utilization of transformer-based models becomes essential in order to effectively capture spatial information across different scales and address the challenges posed by these tasks.

Chapter 5

Pyramid Fine- and Coarse-grained Attentions for Classifying Different Land Cover Types using Compact Polarimetric SAR Data

Land cover classification from RCM CP imagery is important but challenging due to class signature ambiguity issues and speckle noise. This chapter presents a new land cover classification method to improve the learning of discriminative features based on a novel pyramid fine- and coarse-grained self-attentions transformer (PFC transformer). The fine-grained dependency inside of a non-overlapping window and coarse-grained dependencies between non-overlapping windows are explicitly modeled and concatenated using a learnable linear function. This process is repeated in a hierarchical manner. Finally, the output of each stage of the proposed method is spatially reduced and concatenated to take advantage of both low- and high-level features. Two high-resolution (3m) RCM CP SAR scenes are used to evaluate the performance of the proposed method and compare it to other state-of-the-art deep learning methods. The results show that the proposed approach achieves an overall accuracy of 93.63% which was 4.83% higher than the best comparable method, demonstrating the effectiveness of the proposed approach for land cover classification from RCM CP SAR images. This chapter is based on the submitted paper [58].

5.1 Introduction

Land cover classification is essential because it provides valuable information about the Earth's surface and its changes over time which are important for urban planning, natural resource management, and environmental monitoring [13, 14]. Due to the limited data availability, the potential of generating land cover maps using CP SAR data remains largely unexplored. Land cover classification is challenging due to speckle noise [25] and ambiguities associated with backscatter and unique class discrimination [51]. To mitigate this, conventional land cover classification methods increase the number and type of handcrafted features [160]. It is known that pixel-level features and spatially-based texture features have limited capabilities for scene classification [25].

Deep learning (DL) methods provide an advantage over shallow-structured machine learning tools (such as support vector machine [161]) by inherently extracting features [27, 124]. Due to the intrinsic 2-D structure of remote sensing images, CNNs, as a DL approach, are widely used for image processing tasks [162]. While CNNs are able to extract local features, they do not inherently capture long-distance dependency among pixels which is important for land cover classification tasks due to spatial heterogeneity of targets [163]. In contrast, vision transformer methods are capable of capturing long-distance dependencies [164]. As an example, the Vision Transformer (ViT) [165] utilizes the idea of self-attention [166] to enable global receptive field processing of non-overlapping patches.

Despite successful performance on various computer vision tasks [167], ViT has limitations of requiring high computational and memory costs, even for nominally-sized input images and keeping the dimensions of the produced feature maps consistent [168]. To enhance the accuracy and efficiency of ViT in different tasks, several transformer architectures have been introduced [29, 168–170]. These approaches are local-based such as Swin Transformer [29] or global-based such as Pyramid Vision Transformer (PVT) [168]. The local-based approaches divide the input image patch into non-overlapping windows and calculate the self-attention inside of each window. The Swin Transformer uses a shifting window to describe the relationship among windows, which gradually moves the local window's boundaries. However, the window shifting technique lacks optimization for GPU usage and demonstrates inefficient memory utilization [169]. Global approaches such as PVT preserve the global receptive field of ViT but lower the resolution of the key and value feature maps to reduce complexity. However, despite this reduction, the model's complexity is frequently still quadratic in relation to the input image's resolution, posing issues for larger images [169].

Successful classification has been demonstrated by both the local self-attention methods [29, 171] and the global self-attention methods [168, 169]. However, these approaches

impose limitations on the original full self-attention's ability to concurrently capture short- and long-range dependencies [167]. Land cover exhibits high spatial heterogeneity [50]; therefore, capturing both fine-grained and coarse-grained spatial dependencies simultaneously is important because it allows for a comprehensive understanding of the relationships between different pixels in a given feature map. The Focal transformer [167] is designed to integrate fine-grained and different scale coarse-grained spatial dependencies, but to accomplish this task requires a highly complex architecture with accompanying high computing requirements.

In a DL model, the shallow layers primarily focus on capturing low-level and fine features. On the other hand, the deep layers of the model are responsible for extracting deeper, coarse, and semantic features that encapsulate higher-level features, including abstract representations and complex relationships within the data [27]. Consequently, by integrating both low-level and high-level features, the DL model can leverage the complementary nature of these features and achieve a more robust and accurate performance in classification tasks [172].

To the best of our knowledge, there is currently no published research specifically addressing the generation of land cover maps in CP SAR imagery using a self-attention method. As a result, this paper proposes a novel classification method called PFC transformer (Pyramid of Fine- and Coarse-grained attentions transformer), which utilizes a pyramid of window-based vision transformers to measure both fine-grained attention within a window and coarse-grained attention between windows. In summary, this study makes the following contributions in CP SAR land cover classification:

- Our proposed method simultaneously utilizes fine- and coarse-grained spatial dependencies, enabling the method to extract more discriminative and detailed features by capturing spatial relationships at different scales. This attribute effectively addresses spatial heterogeneity present in land covers, ultimately leading to more accurate land cover classification.

- Our proposed method incorporates the outputs of different stages and leverages information across multiple scales, resulting in enhanced accuracy for land cover classification. By addressing the challenges of signature ambiguity, this integration of low- and high-level features improves the accuracy of land cover classification.

- The potential of state-of-the-art (SOTA) DL methods in generating accurate land cover maps using CP SAR data is evaluated and compared with that of the proposed method. This thorough assessment not only advances the understanding of DL techniques in this domain but also provides valuable insights for decision-makers

and researchers aiming to utilize SOTA DL method for land cover classification and monitoring in CP SAR data.

Experiments are based on a pair of high-resolution RCM CP SAR scenes. The proposed PFC transformer surpasses SOTA methods, including Swin, Focal, PVT, Twins [169], CAT [173], SepViT [170], and residual-based CNN (ResCNN) [134], in terms of generating accurate land-cover maps. This chapter is organized as follows. Section 5.2 provides a literature review of land cover classification methods utilizing SAR data. Section 5.3 describes the proposed method, and the study area as well as datasets are introduced in Section 5.4. Section 5.5 presents and analyzes the experimental results, and Section 5.6 provides the conclusions of the study.

5.2 Background

5.2.1 Land cover classification using CP SAR data

Most of the existing land cover classification methods using SAR data are based on QP or DP. There are only a few known published papers on land cover classification using CP SAR data [53,55,174]. Robertson *et al.* [174] utilized hand-crafted features derived from CP SAR data and employed a random forest (RF) classifier for producing crop maps. Nonetheless, the creation of efficient hand-crafted features necessitates expertise in the field and a deep comprehension of the particular domain. Furthermore, the RF classifier does not consider spatial information. Roy *et al.* [53] proposed a MapReduce-based multi-layer perceptron algorithm to distinguish different land cover classes. However, the algorithm did not utilize contextual information, and only numerical results are reported without a classified land cover map, so visual evaluation is not possible. Ghanbari *et al.* [55] proposed a region-based semi-supervised graph network land cover classification using RCM CP SAR data. Despite achieving reliable outcomes, the utilization of hand-crafted features and uncertainty in the homogeneity of generated regions may impact the results. Therefore, it is imperative to focus on designing a feature learning-based land cover classification method for large CP SAR scenes that reduces reliance on hand-crafted features and effectively addresses the issue of signature ambiguity by incorporating spatial information.

5.2.2 Land cover classification using CNNs

CNNs are widely used to generate SAR land cover maps [175]. Zhou *et al.* [176] applied a CNN for QP SAR land cover classification, employing a method that included two convolutional layers and two fully connected layers. Then, several methods for land cover classification based on CNNs were proposed [14, 51, 177–181]. For example, Zhang *et al.* [177] proposed a complex-valued CNN that was tailored to accommodate the arithmetic features of complex data. To extract both spatial- and channel-wise information, Dong *et al.* [178] utilized 3-D convolution. Liu *et al.* [51] considered the statistical distribution of the mid-level features generated by a CNN model to increase the generalization of the model. Although CNNs reached reliable results, they can introduce artifacts along the edges of adjacent patches, leading to the over-smoothing of object boundaries and losing of useful spatial resolution detail [182]. Moreover, despite their proficiency in organizing local features, CNNs encounter challenges in capturing spatial dependencies that extend over long distances [164, 183].

In several recent studies [25, 27, 184–187], fully convolutional networks (FCNs) have been identified as another common approach that exhibits promising land cover results. Wang *et al.* [25] proposed an integration of FCN with sparse and low-rank subspace features network to classify QP SAR images. Li *et al.* [188] suggested the utilization of an FCN with a sliding window technique to alleviate the computational burden and minimize memory usage. Mohammadimanesh *et al.* [27] proposed an FCN network including inception and skip connection to utilize richer contextual information and more detailed information in QP SAR data to classify. Henry *et al.* [185] evaluated the potential of three FCNs in extracting roads from high-resolution SAR images. However, the utilization of FCN methods faces a significant hurdle due to the requirement of whole or dense labeled scenes for their training. The scarcity of labeled SAR data, especially in RCM CP data, makes it infeasible to utilize FCN methods [25]. Given the limitations of CNNs and FCNs in capturing fine- and coarse-grained spatial dependencies and the requirement for dense labeled scenes, it is necessary to explore a method that can effectively capture both levels of spatial dependencies in CP SAR data without relying on whole labeled scenes.

5.2.3 Land cover classification using transformers

Recently, the effectiveness of transformer models in remote sensing applications has captured the attention of remote sensing researchers [175, 183, 189–194]. While several studies have employed transformer models to merge optical and SAR images and leverage the benefits of both data types [188, 195–197], the absence of clear optical images of the same area

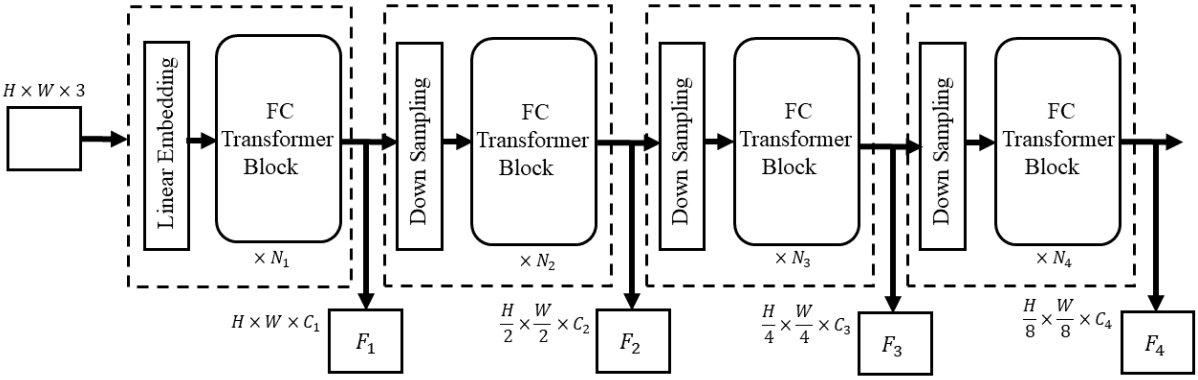


Figure 5.1: Architecture of the PFC-Attention Transformer.

due to cloud cover impedes progress. Other studies have combined CNNs and ViT methods to utilize local and global information for land cover mapping [190, 192, 193]. To integrate the outputs of each branch, various fusion methods have been proposed [192, 195, 198, 199]. However, these methods have certain limitations, such as increased complexity compared to individual models, requiring more time and data for training

To address the limitations discussed, we propose a hierarchical fine- and coarse-grained attentions transformer for land cover classification. Our approach integrates fine and coarse attentions, capturing spatial dependencies, within the same layer using a learnable mechanism. This integration leads to richer information integration. Additionally, our approach leverages a pyramid of low- and high-level features to accommodate varying levels of complexity. By combining these techniques, we aim to overcome the limitations of existing CP SAR land cover approaches and improve accuracy.

5.3 Methodology

Fig 5.1 shows the architecture of the proposed PFC transformer. The proposed approach consists of four stages that produce four feature maps of varying scales. The structure of all stages is similar, comprising of a downsampling layer, except for the stage 1 which includes linear embedding, and N_i times FC block attention. Each part of the architecture is described separately.

5.3.1 Linear Embedding

The linear embedding is a linear transformer that is applied to reduce the spatial size of the image patch and increase the dimension of the raw-valued features into an arbitrary dimension [29, 165]. Since, in this study, the size of the input image patch is not very big, linear embedding is used to increase the dimension of features. Assume that $x_{in} \in R^{H\times W\times C_0}$ is the input image patch where H and W are the spatial dimension and C is the feature dimension, the linear embedding projects x_{in} into $z \in R^{H\times W\times C_1}$.

5.3.2 FC Transformer Block

The main core of the proposed approach is the FC transformer block (see Fig 5.2 (a)). Since the proposed approach is window-based, an input feature map (z) is divided into non-overlapping $M \times M$ windows, and a layer normalization (LN) is applied. Then, by using a linear function, query (Q_f), key (K_f), and value (V_f)$\in R^{(M\times M)\times d}$ matrices are calculated where f stands for fine-grained and d is the depth equals to the feature dimension of z divided by the number of heads [29].

To calculate fine-coarse attention (F_{attn}), similar to the approach employed by the Swin transformer, the self-attention within each window is computed as follows:

$$F_{attn} = softmax(Q_f K_f^T/\sqrt{d} + B_f)V_f \tag{5.1}$$

As described by Liu *et al.* [29], B_f is the learnable relative position bias which its values are taken from $\hat{B}_f \in R^{(2M-1)\times(2M-1)}$. Fig 5.2 (b) shows the structure of the F_{attn}.

In addition to the fine-grained attention, the PFC transformer introduces an approach for calculating the coarse-grained attention (C_{attn}). To compute C_{attn}, a learnable window pooling is applied to reduce the size of K_f and V_f matrices from their original dimensions of $(M \times M) \times d$ to a compact $(1 \times 1) \times d$ representation, called K_c and V_c where c stands for coarse-grained. This reduction in size not only helps to alleviate the complexity and computational costs associated with the model but also enables the consideration of far spatial dependencies. Then, the attention between each fine-grained query matrix, Q_f, and coarse-grained K_c and V_c are calculated as follows:

$$C_{attn} = softmax(Q_f K_c^T/\sqrt{d} + B_c)V_c \tag{5.2}$$

B_c is the relative position bias among fine- and coarse-grained windows; however, since the size of the K_c and V_c are not the same as Q_f, to represent the relative position bias between

them, values in B_c are taken from $\hat{B}_c \in R^{(NW+M-1)\times(NW+M-1)}$ where NW stands for the number of coarse-grained windows [167]. By leveraging this coarse-grained attention mechanism, the model gains the ability to capture broader contextual information and long-range dependencies. The structure of the C_{attn} is depicted in Fig 5.2 (c).

Finally, to take advantage of both fine- and coarse-grained spatial dependencies and utilize them simultaneously, both attentions are concatenated. Nevertheless, this concatenation operation leads to a doubling of the feature dimension. As a result, to restore the number of features to its original value in the input, a projection step becomes necessary. This projection ensures compatibility and coherence in subsequent stages of the computation. The fine- and coarse- attention (FC_{attn}) is computed as:

$$FC_{attn} = Concat(F_{attn}, C_{attn})W_{fc} \tag{5.3}$$

where W_{fc} is the learnable linear projection.

The rest of the FC transformer block is followed by a skip connection with the input feature map, an LN, and a 2-layer multi-layer perceptron (MLP) with GELU nonlinearity in between and again a skip connection, following the same procedure as [29, 165]. In general, the FC transformer block is computed as

$$\alpha = LN(z^{l-1}) \tag{5.4}$$

$$\hat{z}^l = FC_{attn}(F_{attn}(\alpha), C_{attn}(\alpha)) + z^{l-1} \tag{5.5}$$

$$z^l = MLP(LN(\hat{z}^l)) + \hat{z}^l \tag{5.6}$$

in which z^{l-1} is the input feature map from the previous layer.

5.3.3 Downsampling

As the network becomes deeper, reducing the spatial dimensions of the feature maps to produce a hierarchical representation is necessary. Therefore, the downsampling layer which is a convolutional operator compromised of a 2×2 kernel with stride 2 along with an adjustable number of output features is employed to reduce the spatial size of feature maps by a factor of 2. The downsampling reduces the computational cost and allows the network to learn a hierarchical representation of the input.

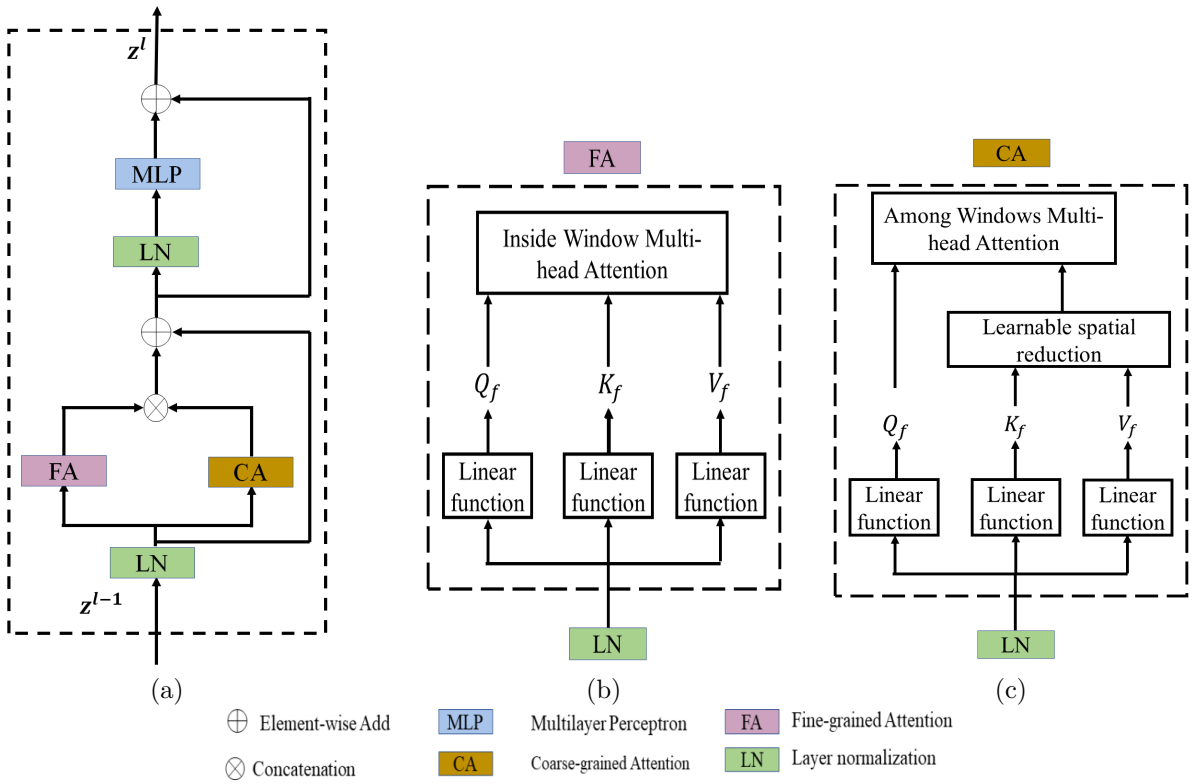


Figure 5.2: (a) Fine- and Coarse-grained Attention Block, (b) Fine-grained Attention, (c) Coarse-grained Attention.

5.3.4 Fusion

Unlike previous methods which only utilized the output of the last stage, our proposed method employs a pyramid of FC transformer blocks' outputs to aggregate information from all stages, allowing for more comprehensive characterization of land cover types (see Fig. 5.1). Given an input image patch of size $H \times W \times 3$, the linear embedding is applied to increase the number of features to C_1. Then, it is passed through an FC transformer block with N_1 layers resulting in F_1 with the shape of $H \times W \times C_1$. Next, F_1 is used as the input of the next stage and this process is repeated to obtain feature maps of F_2, F_3, and F_4. To combine F_1, F_2, F_3, F_4, a learnable linear function is applied to decrease their spatial sizes to that of the final stage's output, which is $H/8 \times W/8$, as follows:

$$F_t = Concat(F_1 W_1, F_2 W_2, F_3 W_3, F_4 W_4) \tag{5.7}$$

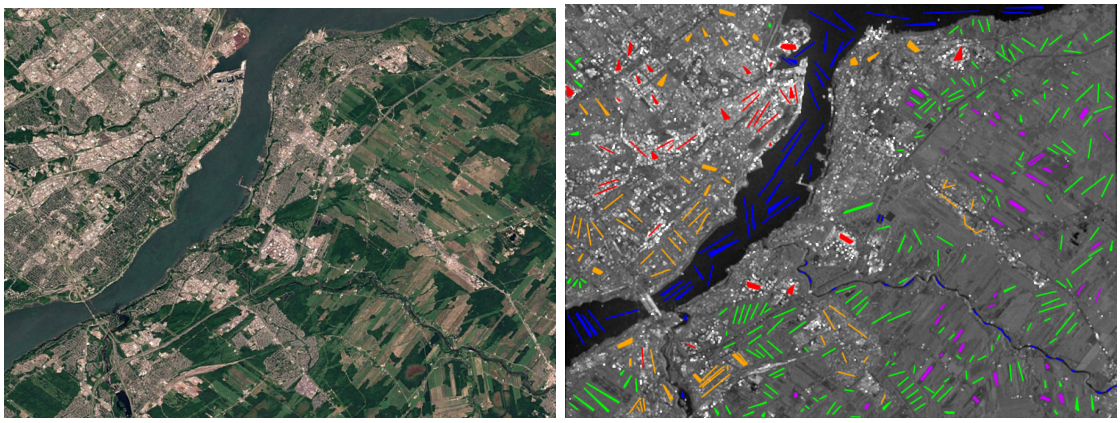


Figure 5.3: (a) Google Earth image of Quebec City scene, (b) the first element of CP coherency matrix along with manually selected samples.

in which W_1, W_2, W_3, and W_4 are convolutional operators with strides of 8, 4, 2, and 1 respectively. The size of F_t is $H/8 \times W/8 \times (C_1 + C_2 + C_3 + C_4)$. Then, a global average pooling layer is applied to F_t followed by a fully connected layer with nodes equal to the number of classes to determine the land cover class.

5.4 Study Area and Dataset

Two very high-resolution (3m) SLC RCM CP SAR scenes with the sampled pixel and line spacing of 1.39 and 2 meters were used to evaluate the performance of the proposed approach and compare it with other approaches. Captured on August 9th, 2022, over Quebec City in Canada, the first scene covers approximately 43 km × 13 km and has a size of 10954 × 8146 pixels, with an incidence angle range of 47.50 to 48.67 degrees. Fig 5.3 (a) presents the Google Earth image of this scene. The second scene, acquired on June 27th, 2020, has a size of 9344×21942 pixels, covering around 43 km×130 km over the city of Ottawa in Canada. Its incidence angle ranges from 38.48 to 39.90 degrees, and its corresponding Google Earth image is shown in Fig 5.4 (a).

The study area has five primary classes: forest, water, two distinct urban areas, and agricultural lands (farms). The urban areas are divided into two groups because some buildings appear bright (Urban 1) while other ones are a mixture of trees and buildings

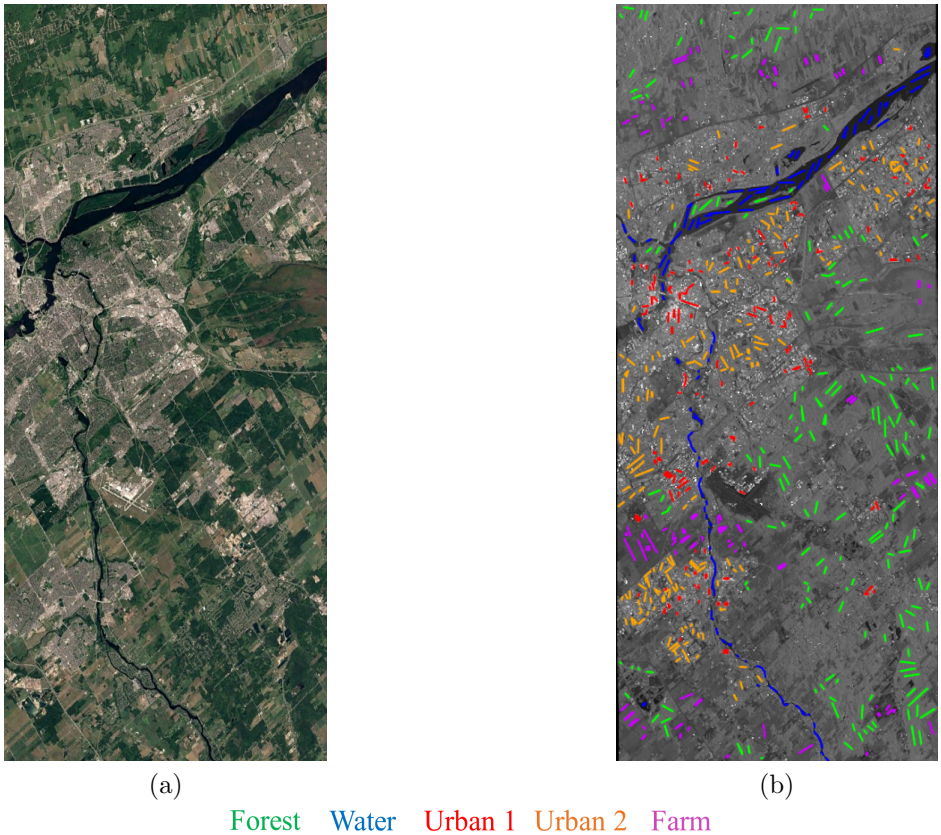


Figure 5.4: (a) Google Earth image of Ottawa scene, (b) the first element of CP coherency matrix along with manually selected samples.

(Urban 2) and their backscattering is not as bright as the first group. The samples were chosen manually by visually examining the SAR scenes and the corresponding Google Earth images.

A 7×7 boxcar filter is applied on both datasets to reduce the impact of speckle noise. Since the images are large and this leads to exceptional computational cost, we reduce this cost by taking a 4×4 non-overlapping block-wise average of the pixels.

Table 5.1: Detailed architecture of the proposed PFC transformer.

	Output	PFC Transformer
Stage 1	32×32×16	Linear Embedding, LN
		$\left\{\begin{matrix}\text{window size} : 4 \times 4 \\ \#\text{heads} : 1\end{matrix}\right\} \times 2$
Stage 2	16×16×32	Downsampling, LN
		$\left\{\begin{matrix}\text{window size} : 4 \times 4 \\ \#\text{heads} : 4\end{matrix}\right\} \times 2$
Stage 3	8×8×64	Downsampling, LN
		$\left\{\begin{matrix}\text{window size} : 4 \times 4 \\ \#\text{heads} : 4\end{matrix}\right\} \times 2$
Stage 4	4×4×128	Downsampling, LN
		$\left\{\begin{matrix}\text{window size} : 4 \times 4 \\ \#\text{heads} : 8\end{matrix}\right\} \times 2$
Global Average	1×1×128	4×4 average pool
Classification	5	128×5 fully connected
Softmax	5	

5.5 Experiments

In this section, the performance of the proposed method in classifying land type covers is discussed and compared to that of the SOTA methods. To assess the efficacy of combining features with different levels, the proposed method was applied both with (PFC transformer) and without (FC transformer) using the pyramid of features. Table 5.1 indicates the structure of the proposed method. It includes 4 stages where the FC transformer block is repeated twice in each stage. The number of feature maps in each stage is set to 16, 32, 64, and 128, respectively. The size of non-overlapping windows is set to 4×4, and the number of heads for each stage is 1, 4, 4, and 8, respectively.

The performance of the methods was examined using several metrics, namely, overall accuracy (OA), kappa coefficient (κ), f-1 scores of each class ($F1$), and averaged f-1 score ($F1_{avg}$). OA is determined by dividing the number of correctly classified test samples by

Table 5.2: The number of training and testing pixels for each class selected from the Quebec and Ottawa scenes, respectively.

Class	# of train	# of test
Forest	15381	11690
Water	14853	8093
Urban 1	12032	11263
Urban 2	15098	10206
Farm	10773	20022

the total number of test samples. κ measures the level of agreement between the test samples and the final labeled map [27]. $F1$ is a harmonic mean of precision and recall, which is particularly useful for imbalanced classes [27]. The highest and lowest possible values of $F1$ are 1 and 0 and it is calculated as:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{5.8}$$

where

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{5.9}$$

and

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{5.10}$$

Precision metric measures the accuracy of detected pixels in each class while *Recall* indicates the number of true pixels identified in each class.

5.5.1 Training and Testing

In this study, the labeled pixels of the Quebec scene were used for training the models that were evaluated using the labeled pixels chosen from the Ottawa scene. Moreover, to better evaluate the performance of the models, three different regions have been selected and shown in Fig 5.5. Regions A and B show agricultural, forest and urban areas, while Region C includes forest and agricultural areas.

Table 5.2 represents the number of training and testing samples. The training samples were used to standardize the Quebec and Ottawa scenes. To train the models, patches of

Figure 5.5: The Google Earth image of the test scene with three regions of interest along with their corresponding $|S_{RH}|^2$ scene. Regions A and B primarily consist of urban, farm and forest classes, while Region C displays both forest and farm classes.

size 32×32×3 were extracted around each labeled pixel, where 3 represents the absolute value of the coherency matrix elements in (2.5). In addition, the models were trained using ADAMW optimization [200] with the learning rate, weight decay, and beta parameters set to 1e-5, 0.05, 0.9, and 0.999 as well as the batch size and training epochs are 32 and 100, respectively. In the training step, 80% of the training samples were utilized to adjust the model's weight values by minimizing the multi-class cross-entropy lost function [153], while the remaining 20% were used for validation purposes. The weight values of the model that achieved the highest validation accuracy were selected.

5.5.2 Results

Fig. 5.6 shows the results obtained by the different models along with their OA. Due to resizing the images to fit the page, finer details present in the original images are not

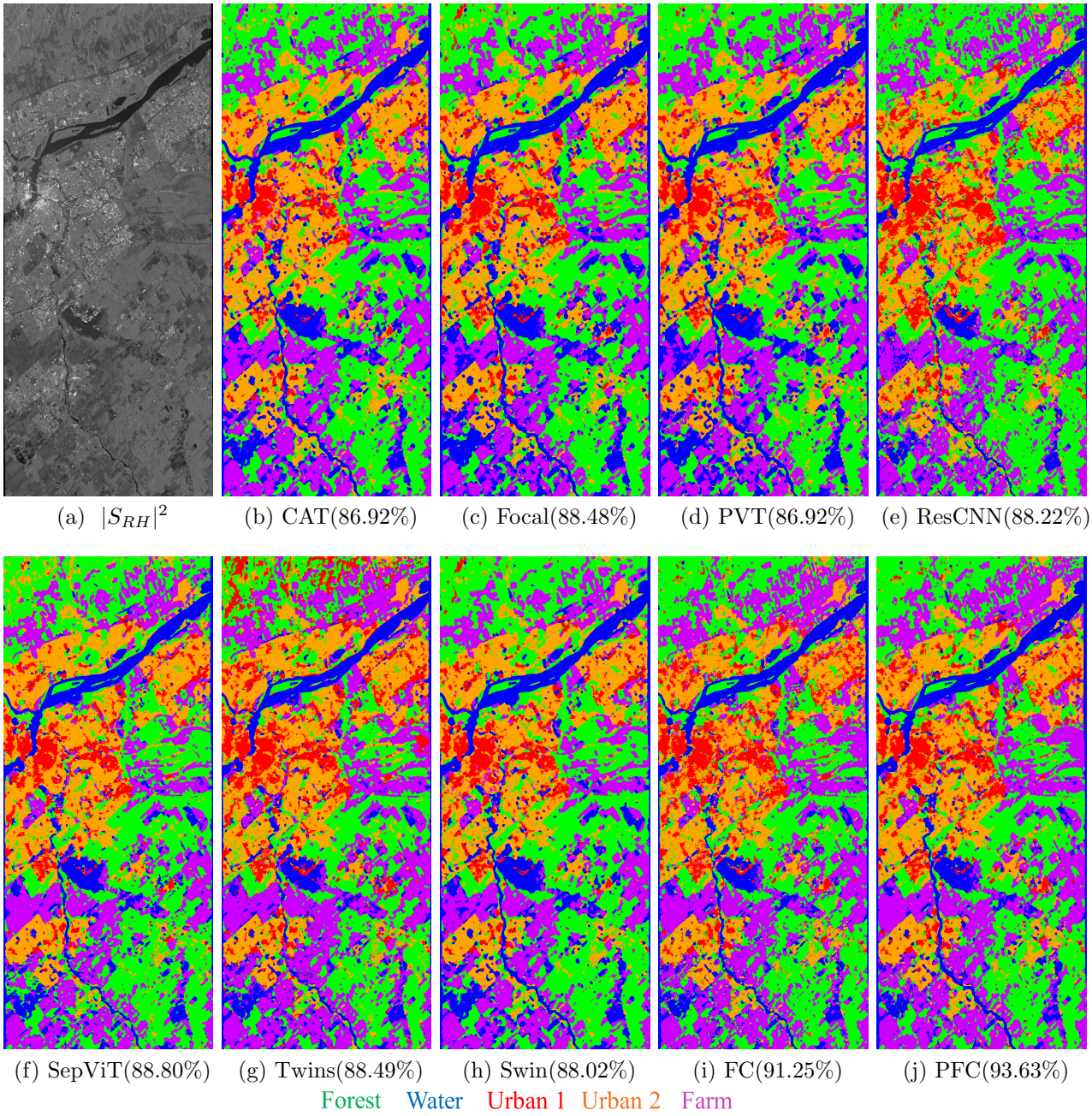
Table 5.3: Assessment of the results obtained by the different models by using overall accuracy (OA), kappa coefficient (κ), averaged f-1 score ($F1_{avg}$), and f-1 score of each class. the **bold** numbers indicate the highest accurate results.

Name	$OA(\%)$	κ	$F1_{avg}$	Forest	Water	Urban1	Urban2	Farm
CAT [173]	86.92	0.8343	0.8696	0.9116	0.7723	0.8843	0.9234	0.8574
Focal [167]	88.48	0.8544	0.8852	0.9248	0.7812	0.9208	**0.9278**	0.8715
PVT [168]	86.92	0.8351	0.8694	0.9218	0.7596	0.8955	0.9138	0.8561
ResCNN [134]	88.22	0.8500	0.8780	0.8835	0.8246	**0.9215**	0.8820	0.8984
SepViT [170]	88.80	0.8579	0.8850	0.9049	0.8219	0.9046	0.8971	0.8970
Twins [169]	88.49	0.8538	0.8788	0.9181	0.8153	0.8531	0.8931	0.9144
Swin [29]	87.14	0.8372	0.8707	0.9049	0.7881	0.9076	0.8866	0.8661
FC	91.25	0.8885	0.9054	0.9346	0.8564	0.9087	0.8844	0.9428
PFC	**93.63**	**0.9185**	**0.9285**	**0.9491**	**0.8864**	0.9191	0.9179	**0.9701**

apparent. Upon visual inspection of the outputs, the CAT, Focal, PVT, Swin, and ResCNN models appear to overestimate the water class in the lower portion of the scene. Twins misclassifies many forest and farm samples into Urban 1 class in the upper part of the scene. SepViT and Twins exhibited poor detection of the river in the middle of the scene and it is narrow than that detected by the other models as well as the proposed ones.

The FC and PFC transformer models have a higher accuracy and improved spatial representation in specifying the type of land covers than the SOTA ones. This is because the proposed models, unlike the other approaches, utilize close and far dependency among pixels simultaneously and combine different feature levels resulting in reducing the rate of misclassification.

As shown in Table 5.3, we compared the quantitative results obtained by the proposed models to those of the SOTA ones. The CAT and PVT models were found to have the lowest OA among the SOTA ones, with both achieving of 86.92%. In contrast, the SepViT model achieved a reliable overall accuracy of 88.80%. While the Swin, CAT, and PVT models showed comparable κ and $F1_{avg}$, the Focal, ResCNN, SepViT, and Twins models obtained higher accuracies, albeit still lower than those achieved by the FC and PFC transformer models.



(a) $|S_{RH}|^2$ (b) CAT(86.92%) (c) Focal(88.48%) (d) PVT(86.92%) (e) ResCNN(88.22%)

(f) SepViT(88.80%) (g) Twins(88.49%) (h) Swin(88.02%) (i) FC(91.25%) (j) PFC(93.63%)

Figure 5.6: (a) shows the $|S_{RH}|^2$ image of the test scene and (b)-(j) are the results obtained by the SOTA models and the proposed ones along with their OA.

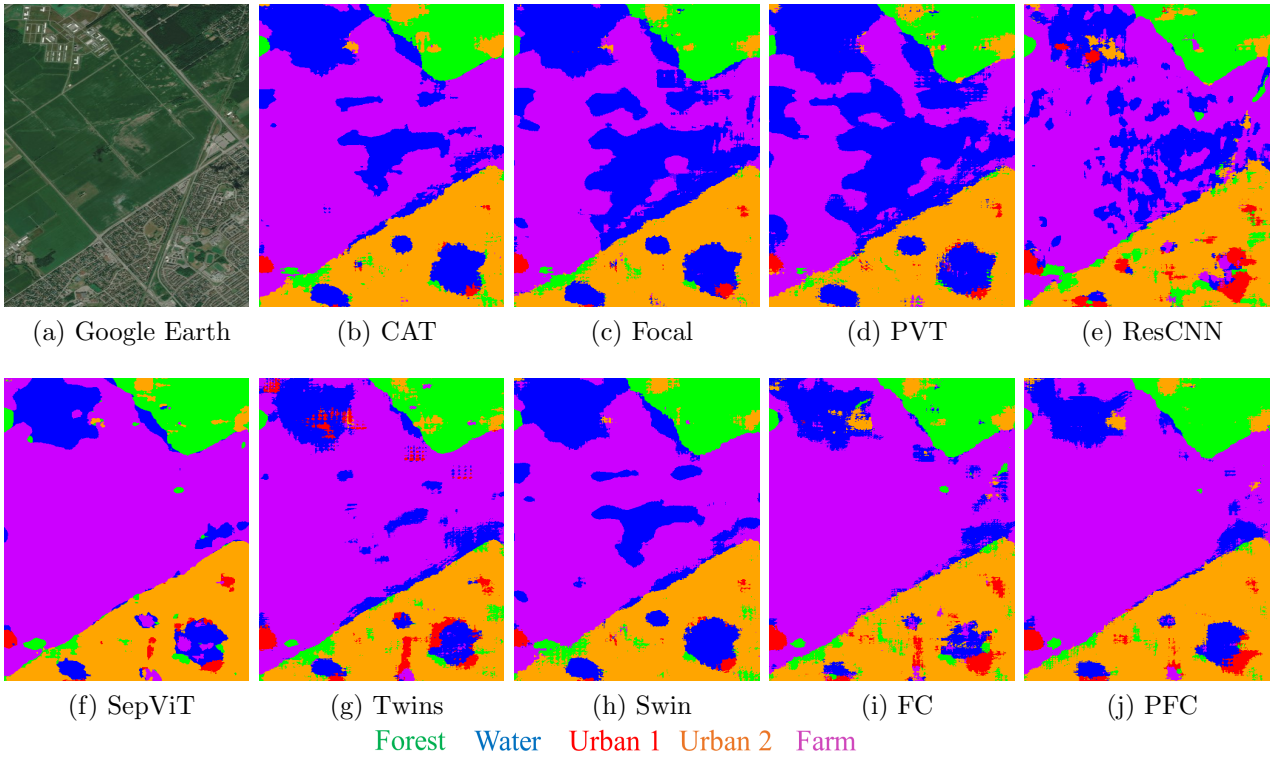


Figure 5.7: (a) shows the Google Earth image of Region A including urban, farm, and forest classes. (b)-(j) are the results obtained by each model.

The proposed FC transformer model achieved an OA of 91.25%, which is about 3-4% higher than those achieved by the SepViT and CAT models. The higher values of κ and $F1_{avg}$ obtained by the FC transformer model provide additional evidence of the effectiveness of the proposed attention mechanism in improving the accuracy of generating land cover type maps. These findings identify that the SOTA models have limitations that prevent them from achieving the same level of performance as the FC transformer.

When comparing the performance of the FC and PFC transformer models, we found that the PFC transformer outperformed the former, with a higher accuracy. By fusion of the different feature levels in a learnable manner, the OA value reached 2% higher. Moreover, the higher values of κ and $F1_{avg}$ obtained by the PFC transformer suggest that leveraging the outputs of all stages can lead to improved accuracy of the balanced and imbalanced classes [160].

The FC and PFC transformers yield higher $F1$ for the forest, water, and farm classes

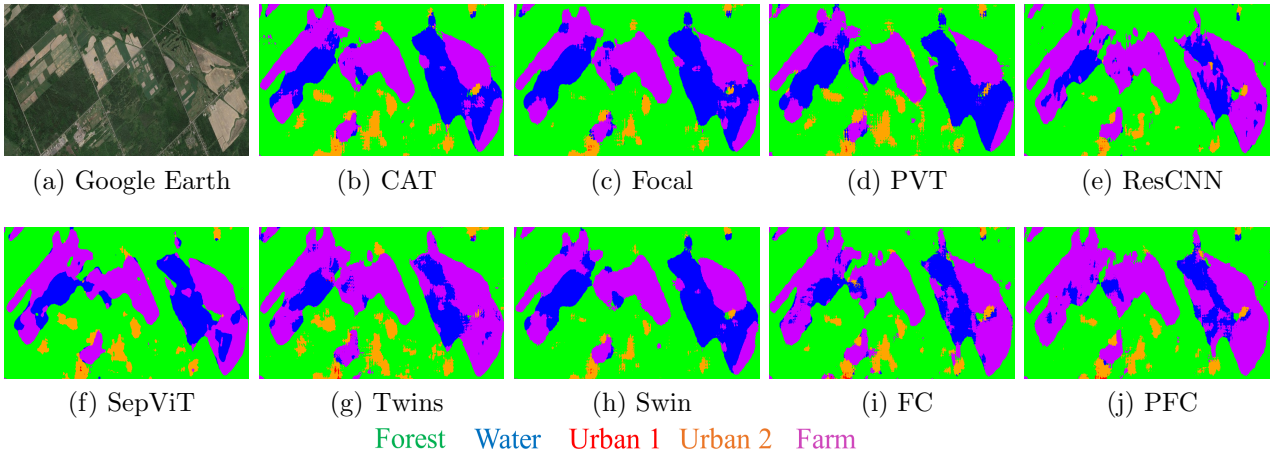


Figure 5.8: (a) displays a Google Earth image of Region B, which includes agricultural lands, forests, and a few buildings. (b)-(j) are the results obtained by each model.

than the SOTA models, demonstrating the significance of fine- and coarse-grained dependencies among pixels and the benefits of utilizing features at different levels. The ResCNN and Focal models achieved slightly higher $F1$ for the urban classes compared to the proposed models, but the difference is negligible.

Fig. 5.7 shows the outputs of the models on Region A which is a mixture of buildings, forest, and agricultural areas. Notably, the CAT, Focal, PVT, ResCNN, and Swin models exhibited a higher rate of misclassifying water in this region, while the SepViT, Twins, and proposed models yielded more accurate outcomes. The output of the models for Region B is shown in Fig. 5.8. Among the SOTA models, the CAT, Focal, PVT, SepViT, and Swin models misclassified a significant portion of the agricultural lands as water class while the ResCNN and Twins performed better. Moreover, the FC transformer exhibited performance over ResCNN and Twins, but the PFC transformer achieved the best classification performance in Region B. Using fine- and coarse-grained spatial information decreased the rate of water misclassification in particular for the left agricultural land. By adding the pyramid of low- and high-level features to the FC transformer, the rate of misclassification was reduced significantly. This is because the integration of different level features enables the model to capture a wide range of features across different scales. Fig. 5.9 shows the output of the models on Region C, which includes forest and farm classes. All models, except the proposed ones, had a high rate of misclassifying agricultural areas as forests. The proposed models exhibited significantly better classification performance for forests.

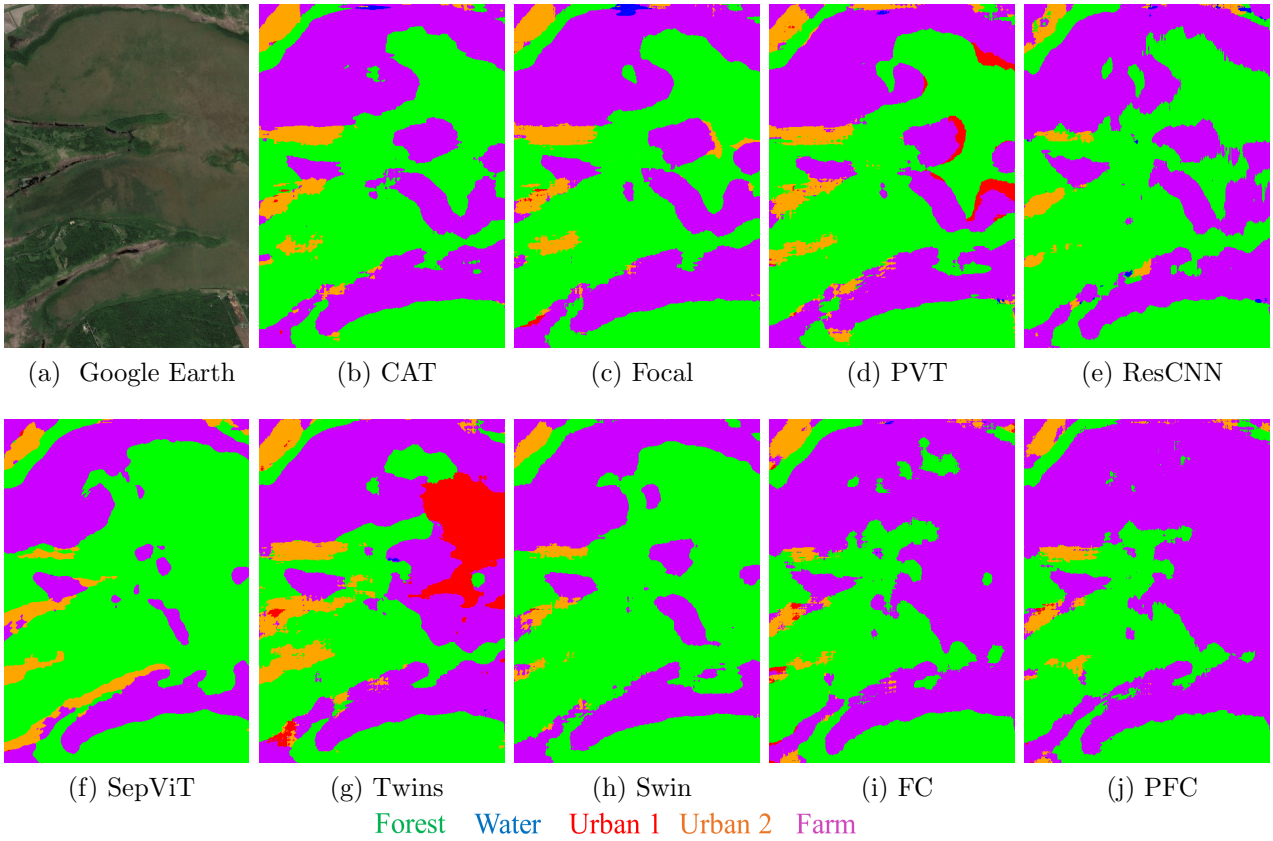


Figure 5.9: (a) shows the Google Earth image of Region C including forest and farm classes. (b)-(j) are the results obtained by each model.

5.6 Conclusion

A new transformer approach was introduced in this chapter for automated generation of land cover maps using high-resolution SAR scenes. To the best of our knowledge, this is the first study that leverages spatial attention information in CP SAR data for land type classification. The proposed attention mechanism captures both fair- and coarse-grained dependencies among pixels within a feature map, resulting in richer information. This attribute endows the method with the ability to consider the spatial relationship among the pixels, resulting in more accurate outputs. The qualitative and quantitative comparison of the results obtained by the proposed attention method and the well-known SOTA methods confirms the efficiency of using short and long dependencies simultaneously

in increasing the accuracy of the generated land cover maps.

Furthermore, we take into account the outputs from all stages and exploit the information across various scales to utilize more detailed information. By utilizing this fusion methodology, the proposed method's capability to identify land cover types is enhanced. Comparing the outputs of the proposed method with and without feature fusion reveals that features at various levels contain valuable information that must be taken into consideration.

The limited availability of RCM data has led to a shortage of annotated CP scenes. As training deep learning methods demand a large number of samples, it is essential to consider semi-supervised techniques in studies. The proposed method can potentially be applied for dense semantic segmentation purposes by increasing the availability of RCM CP SAR scenes and ground truth samples in the future.

Chapter 6

Conclusion and Future Work

The primary objective of this thesis was to develop classification algorithms specifically designed for CP SAR data. The emergence of the RCM, which provides access to CP data, highlights the significance of automating CP classification methods for a wide range of applications, particularly in environmental monitoring tasks such as oil spill detection, sea ice mapping, and land cover mapping.

To address the challenges posed by signal ambiguities influenced by factors like weather conditions (e.g., wind speed) and satellite infrastructure (e.g., incidence angle or speckle noise), this thesis proposes three classification algorithms tailored to CP SAR data, utilizing spatial modeling techniques.

In Chapter 3, the proposed CRF model incorporates the statistical characteristics of CP SAR data and effectively models spatial interactions between pixels by measuring similarities in features and labels. The CRF model demonstrates superior accuracy compared to the baseline MRF and CRF models.

Chapter 4 introduces a developed sea ice classification methodology that eliminates the need for hand-crafted features and instead employs feature learning methods to harness the spatial information inherent in CP SAR data. Additionally, the sea ice maps are integrated with regions obtained through a tailored CP SAR segmentation method., resulting in improved performance compared to previous CP sea ice classification techniques.

Finally, a novel transformer classification approach was proposed to capture both fine and coarse spatial dependencies between pixels and leverage information at different scales. This approach, discussed in Chapter 5, addressed the challenge of high heterogeneity in target distribution and signature ambiguity commonly encountered in land cover mapping

tasks. By effectively modeling spatial dependencies and incorporating multi-scale features, the proposed transformer classification approach significantly improved the accuracy of land cover type detection.

The experimental results showcased the superiority of the proposed methods over commonly used and state-of-the-art approaches when evaluated on diverse datasets. The contributions and outcomes of each chapter were summarized, highlighting the advancements achieved in CP SAR data classification for various applications.

In conclusion, this thesis has made significant contributions to the field of CP SAR data classification by introducing tailored algorithms that effectively address signal ambiguities and leverage spatial modeling techniques. The proposed methods have demonstrated superior performance and have the potential to enhance the accuracy and reliability of CP SAR data analysis for environmental monitoring tasks and beyond.

6.1 Summary of Contributions and Results

This thesis improved CP SAR classifications in three key applications, i.e, oil spill candidate detection, sea ice mapping, and land cover mapping by modeling spatial information. This thesis has the following main contributions.

- In Chapter 3, a CRF model was designed to enhance the detection of oil spill candidates. This CRF model stands out by leveraging the statistical properties of CP SAR data. Moreover, it incorporates the similarities between labels and features of pixels to effectively capture and integrate spatial information into the detection process. By combining these properties, the proposed CRF model offers a comprehensive and tailored solution for improving oil spill candidate detection in CP SAR data. The obtained results clearly demonstrated the superior performance of the proposed CRF model compared to traditional MRF and CRF models in accurately identifying oil spill candidates. These findings highlight the criticality of leveraging appropriate statistical characteristics of the data and integrating spatial information within the modeling process.

- In Chapter 4, a methodology for CP sea ice classification was developed as another contribution of this thesis. The methodology involves multiple steps to enhance the accuracy of sea ice mapping. First, a full scene segmentation is performed to generate class-homogeneous regions. Depending on the type of SAR data, either IRGS (in the case of DP or CP intensities) or CP-IRGS (in the case of CP coherence matrix),

or PolarIRGS (in the case of RQP and QP coherency matrices) are performed to generate regions. Next, a ResCNN model is employed to obtain pixel-level sea ice maps by effectively learning sea ice features and spatial patterns from CP SAR data. Finally, the regions and pixel-level sea ice maps are integrated using a majority voting approach, enabling the assignment of accurate ice labels to each region. This integration process significantly enhances the resulting sea ice map, contributing to improved sea ice classification performance. Based on the experimental results, the proposed methodology produces sea ice maps with higher accuracy compared to DP and RQP methods, while demonstrating comparable performance to the QP approach. Furthermore, the ResCNN classifier reached a higher overall classification accuracy (88.23%) compared to the baseline methods (87.02% and 87.60%), without the need for polarimetric and GLCM features. These findings provide compelling evidence that feature learning classifiers have the potential to enhance the accuracy of sea ice maps beyond the capabilities of traditional machine learning methods like SVM.

- In Chapter 5, as the third contribution of this thesis, a novel transformer approach is proposed for the automated generation of high-resolution land cover maps using SAR scenes. The proposed attention mechanism effectively captures both fine- and coarse-grained dependencies among pixels within a feature map, thereby incorporating richer spatial information. This attribute empowers the method to consider the complex spatial relationships among pixels, resulting in more accurate outputs. Through qualitative and quantitative comparisons with well-known SOTA methods, the results validate the effectiveness of the proposed attention method in enhancing the accuracy of generated land cover maps, particularly through its ability to handle long dependencies.

6.2 Future Work

In this thesis, the utilization of spatial information in CP SAR imagery was shown to be advantageous. Although the proposed methods achieved superior results, there are several areas that warrant further investigation and can be explored in future work.

- Semi-Supervised Methods: Given the limited availability of annotated CP SAR data, exploring and developing semi-supervised learning methods can be a promising direction for future research. These techniques can leverage the available labeled data

along with a larger pool of unlabeled data to improve the performance of classification and mapping algorithms. Investigating techniques such as active learning or self-training can help effectively utilize the limited annotated data and expand the training set.

- Region-based Techniques: Considering the time-consuming nature of training deep learning methods, exploring region-based techniques can be beneficial. These approaches focus on extracting relevant features and modeling spatial relationships within segmented regions, reducing the computational burden and enhancing the efficiency of classification algorithms.

- Availability of Annotated CP SAR Data: With the expectation of more available annotated CP SAR data in the future, there is an opportunity to explore advanced techniques such as semantic segmentation. Leveraging larger annotated datasets can facilitate the training of deep learning models specifically designed for semantic segmentation tasks.

References

[1] J. R. Jensen, *Remote sensing of the environment: An earth resource perspective 2/e*. Pearson Education India, 2009.

[2] D. Flett, Y. Crevier, and R. Girard, "The RADARSAT constellation mission: Meeting the government of Canada's needs and requirements," in *2009 IEEE International Geoscience and Remote Sensing Symposium*, vol. 2. IEEE, 2009, pp. II–910.

[3] C. S. Agency, "RADARSAT constellation," 2019. [Online]. Available: https://www.asc-csa.gc.ca/eng/satellites/radarsat/technical-features/characteristics.asp

[4] M. M. Espeseth, C. Brekke, and S. N. Anfinsen, "Hybrid-polarity and reconstruction methods for sea ice with L- and C-band SAR," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 3, pp. 467–471, 2016.

[5] R. K. Raney, "Comparing compact and quadrature polarimetric SAR performance," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 6, pp. 861–864, 2016.

[6] R. K. Panigrahi and A. K. Mishra, "Comparison of hybrid-pol with quad-pol scheme based on polarimetric information content," *International journal of remote sensing*, vol. 33, no. 11, pp. 3531–3541, 2012.

[7] V. Kumar, D. Mandal, A. Bhattacharya, and Y. Rao, "Crop characterization using an improved scattering power decomposition technique for compact polarimetric SAR data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 88, p. 102052, 2020.

[8] F. Charbonneau, B. Brisco, R. Raney, H. McNairn, C. Liu, P. Vachon, J. Shang, R. DeAbreu, C. Champagne, A. Merzouki *et al.*, "Compact polarimetry overview and applications assessment," *Canadian Journal of Remote Sensing*, vol. 36, no. sup2, pp. S298–S315, 2010.

[9] R. K. Raney, "Hybrid-polarity SAR architecture," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 11, pp. 3397–3404, 2007.

[10] 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.

[11] L. P. Bobylev and M. W. Miles, "Sea ice in the Arctic paleoenvironments," *Sea Ice in the Arctic: Past, Present and Future*, pp. 9–56, 2020.

[12] D. G. Barber, "Microwave remote sensing, sea ice and Arctic climate," *Phys. Can*, vol. 61, pp. 105–111, 2005.

[13] Z. Qi, A. G.-O. Yeh, X. Li, and Z. Lin, "A novel algorithm for land use and land cover classification using RADARSAT-2 polarimetric SAR data," *Remote Sensing of Environment*, vol. 118, pp. 21–39, 2012.

[14] X. Li, L. Lei, Y. Sun, M. Li, and G. Kuang, "Multimodal bilinear fusion network with second-order attention-based channel selection for land cover classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 1011–1026, 2020.

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

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] M. S. Mahmud, T. Geldsetzer, S. E. 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.

[21] 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.

[22] 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.

[23] 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*, vol. 14, pp. 5831–5841, 2021.

[24] C. M. Gurney, J. R. Townshend *et al.*, "The use of contextual information in the classification of remotely sensed data," *Photogrammetric Engineering and Remote Sensing*, vol. 49, no. 1, pp. 55–64, 1983.

[25] Y. Wang, C. He, X. Liu, and M. Liao, "A hierarchical fully convolutional network integrated with sparse and low-rank subspace representations for PolSAR imagery classification," *Remote Sensing*, vol. 10, no. 2, p. 342, 2018.

[26] P. Zhong and R. Wang, "A multiple conditional random fields ensemble model for urban area detection in remote sensing optical images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 12, pp. 3978–3988, 2007.

[27] F. Mohammadimanesh, B. Salehi, M. Mahdianpari, E. Gill, and M. Molinier, "A new fully convolutional neural network for semantic segmentation of polarimetric SAR imagery in complex land cover ecosystem," *ISPRS journal of photogrammetry and remote sensing*, vol. 151, pp. 223–236, 2019.

[28] B. Ai, Z. Wen, Z. Wang, R. Wang, D. Su, C. Li, and F. Yang, "Convolutional neural network to retrieve water depth in marine shallow water area from remote sensing images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 2888–2898, 2020.

[29] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 10 012–10 022.

[30] A.-B. Salberg, Ø. Rudjord, and A. H. S. Solberg, "Oil spill detection in hybrid-polarimetric SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 10, pp. 6521–6533, 2014.

[31] F. Nunziata, M. Migliaccio, and X. Li, "Sea oil slick observation using hybrid-polarity SAR architecture," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 2, pp. 426–440, 2014.

[32] H. Li, W. Perrie, and J. Wu, "Retrieval of oil–water mixture ratio at ocean surface using compact polarimetry synthetic aperture radar," *Remote Sensing*, vol. 11, no. 7, p. 816, 2019.

[33] Y. Li, Y. Zhang, J. Chen, and H. Zhang, "Improved compact polarimetric SAR quad-pol reconstruction algorithm for oil spill detection," *IEEE geoscience and remote sensing letters*, vol. 11, no. 6, pp. 1139–1142, 2013.

[34] R. Shirvany, M. Chabert, and J.-Y. Tourneret, "Ship and oil-spill detection using the degree of polarization in linear and hybrid/compact dual-pol SAR," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 3, pp. 885–892, 2012.

[35] M. M. Espeseth, S. Skrunes, C. E. Jones, C. Brekke, B. Holt, and A. P. Doulgeris, "Analysis of evolving oil spills in full-polarimetric and hybrid-polarity SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 4190–4210, 2017.

[36] Y. Li, H. Lin, Y. Zhang, and J. Chen, "Comparisons of circular transmit and linear receive compact polarimetric SAR features for oil slicks discrimination," *Journal of Sensors*, vol. 2015, 2015.

[37] A. Stumpf and N. Kerle, "Combining random forests and object-oriented analysis for landslide mapping from very high resolution imagery," *Procedia Environmental Sciences*, vol. 3, pp. 123–129, 2011.

[38] Y. Wu, K. Ji, W. Yu, and Y. Su, "Region-based classification of polarimetric SAR images using Wishart MRF," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 4, pp. 668–672, 2008.

[39] S. Singha and R. Ressel, "Arctic sea ice characterization using RISAT-1 compact-pol SAR imagery and feature evaluation: A case study over Northeast Greenland," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 8, pp. 3504–3514, 2017.

[40] M. Dabboor and T. Geldsetzer, "Towards sea ice classification using simulated RADARSAT Constellation Mission compact polarimetric SAR imagery," *Remote Sensing of Environment*, vol. 140, pp. 189–195, 2014.

[41] T. Geldsetzer, M. Arkett, T. Zagon, F. Charbonneau, J. J. Yackel, and R. K. Scharien, "All-season compact-polarimetry C-band SAR observations of sea ice," *Canadian Journal of Remote Sensing*, vol. 41, no. 5, pp. 485–504, 2015.

[42] H. Li and W. Perrie, "Sea ice characterization and classification using hybrid polarimetry SAR," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 11, pp. 4998–5010, 2016.

[43] 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.

[44] X. Zhang, J. Zhang, M. Liu, and J. Meng, "Assessment of C-band compact polarimetry SAR for sea ice classification," *Acta Oceanologica Sinica*, vol. 35, no. 5, pp. 79–88, 2016.

[45] 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 Sensing*, vol. 10, no. 4, p. 594, 2018.

[46] S. Nasonova, R. K. Scharien, T. Geldsetzer, S. E. Howell, and D. Power, "Optimal compact polarimetric parameters and texture features for discriminating sea ice types during winter and advanced melt," *Canadian Journal of Remote Sensing*, vol. 44, no. 4, pp. 390–411, 2018.

[47] X. Sun, X. Zhang, W. Huang, Z. Han, X. Lyu, and P. Ren, "Sea ice classification using mutually guided contexts," *IEEE Transactions on Geoscience and Remote Sensing*, 2023.

[48] 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.

[49] 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, 2013.

[50] H. Xing, L. Zhu, Y. Feng, W. Wang, D. Hou, F. Meng, and Y. Ni, "An adaptive change threshold selection method based on land cover posterior probability and spatial neighborhood information," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 11 608–11 621, 2021.

[51] X. Liu, C. He, Q. Zhang, and M. Liao, "Statistical convolutional neural network for land-cover classification from SAR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 9, pp. 1548–1552, 2019.

[52] Z. Ren, B. Hou, Z. Wen, and L. Jiao, "Patch-sorted deep feature learning for high resolution SAR image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 9, pp. 3113–3126, 2018.

[53] S. Roy, A. Das, and S. N. Omkar, "A distributed land cover classification of FP and CP SAR observation using MapReduce-based multi-layer perceptron algorithm over the Mumbai mangrove region of India," *International Journal of Remote Sensing*, vol. 44, no. 5, pp. 1510–1532, 2023.

[54] B. Brisco, K. Li, B. Tedford, F. Charbonneau, S. Yun, and K. Murnaghan, "Compact polarimetry assessment for rice and wetland mapping," *International Journal of Remote Sensing*, vol. 34, no. 6, pp. 1949–1964, 2013.

[55] M. Ghanbari, L. Xu, and D. A. Clausi, "Local and global spatial information for land cover semi-supervised classification of complex polarimetric SAR data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023.

[56] S. Taleghanidoozdoozan, L. Xu, and D. A. Clausi, "Oil spill candidate detection using a conditional random field model on simulated compact polarimetric imagery," *Canadian Journal of Remote Sensing*, vol. 48, no. 3, pp. 425–440, 2022.

[57] S. Taleghanidoozdoozan, L. Xu, and D. A. Clausi, "Region-based sea ice mapping using compact polarimetric synthetic aperture radar imagery with learned features and contextual information," *Remote Sensing*, vol. 15, no. 12, p. 3199, 2023.

[58] S. Taleghanidoozdoozan, L. Xu, and D. A. Clausi, "Pyramid fine and coarse attentions for land cover classification from compact polarimetric SAR imagery," *Submitted to IEEE Transactions on Geoscience and Remote Sensing*, 2023.

[59] K. Irwin, A. Braun, G. Fotopoulos, A. Roth, and B. Wessel, "Assessing single-polarization and dual-polarization TerraSAR-X data for surface water monitoring," *Remote Sensing*, vol. 10, no. 6, p. 949, 2018.

[60] J.-S. Lee and E. Pottier, *Polarimetric radar imaging: from basics to applications*. CRC press, 2017.

[61] 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.

[62] M. Jafari, Y. Maghsoudi, and M. J. V. Zoej, "A new method for land cover characterization and classification of polarimetric SAR data using polarimetric signatures," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 7, pp. 3595–3607, 2015.

[63] S. Cloude, *Polarisation: applications in remote sensing*. OUP Oxford, 2009.

[64] M. Denbina and M. J. Collins, "Iceberg detection using compact polarimetric synthetic aperture radar," *Atmosphere-Ocean*, vol. 50, no. 4, pp. 437–446, 2012.

[65] 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.

[66] X. Deng, C. López-Martínez, J. Chen, and P. Han, "Statistical modeling of polarimetric SAR data: A survey and challenges," *Remote Sensing*, vol. 9, no. 4, p. 348, 2017.

[67] F. Edition, A. Papoulis, and S. U. Pillai, "Probability, random variables, and stochastic processes," 2002.

[68] C. Oliver and S. Quegan, *Understanding synthetic aperture radar images*. SciTech Publishing, 2004.

[69] R. Dekker, "Speckle filtering in satellite SAR change detection imagery," *International Journal of Remote Sensing*, vol. 19, no. 6, pp. 1133–1146, 1998.

[70] C. Walck *et al.*, "Hand-book on statistical distributions for experimentalists," *University of Stockholm*, vol. 10, 2007.

[71] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3795–3809, 2009.

[72] J.-S. Lee, M. R. Grunes, and R. Kwok, "Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution," *International Journal of Remote Sensing*, vol. 15, no. 11, pp. 2299–2311, 1994.

[73] J.-S. Lee and E. Pottier, *Polarimetric radar imaging: from basics to applications*. CRC press, 2009.

[74] J.-C. Souyris, P. Imbo, R. Fjortoft, S. Mingot, and J.-S. Lee, "Compact polarimetry based on symmetry properties of geophysical media: The $\pi/4$ mode," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 634–646, 2005.

[75] M. E. Nord, T. L. Ainsworth, J.-S. Lee, and N. J. Stacy, "Comparison of compact polarimetric synthetic aperture radar modes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 1, pp. 174–188, 2008.

[76] S. Nghiem, S. Yueh, R. Kwok, and F. Li, "Symmetry properties in polarimetric remote sensing," *Radio Science*, vol. 27, no. 05, pp. 693–711, 1992.

[77] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 48, no. 3, pp. 259–279, 1986.

[78] 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.

[79] V. Kolmogorov and R. Zabin, "What energy functions can be minimized via graph cuts?" *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 2, pp. 147–159, 2004.

[80] B. Fiscella, A. Giancaspro, F. Nirchio, P. Pavese, and P. Trivero, "Oil spill detection using marine SAR images," *International Journal of Remote Sensing*, vol. 21, no. 18, pp. 3561–3566, 2000.

[81] Y. Shu, J. Li, H. Yousif, and G. Gomes, "Dark-spot detection from SAR intensity imagery with spatial density thresholding for oil-spill monitoring," *Remote Sensing of Environment*, vol. 114, no. 9, pp. 2026–2035, 2010.

[82] C. Brekke and A. H. Solberg, "Oil spill detection by satellite remote sensing," *Remote sensing of environment*, vol. 95, no. 1, pp. 1–13, 2005.

[83] L. De Laurentiis, C. E. Jones, B. Holt, G. Schiavon, and F. Del Frate, "Deep learning for mineral and biogenic oil slick classification with airborne synthetic aperture radar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 10, pp. 8455–8469, 2020.

[84] A. S. Solberg, G. Storvik, R. Solberg, and E. Volden, "Automatic detection of oil spills in ERS SAR images," *IEEE Transactions on geoscience and remote sensing*, vol. 37, no. 4, pp. 1916–1924, 1999.

[85] J.-S. Lee, M. R. Grunes, and G. De Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Transactions on Geoscience and remote sensing*, vol. 37, no. 5, pp. 2363–2373, 1999.

[86] S. M. Ahmed, F. A. E. Eldin, and A. M. Tarek, "Speckle noise reduction in SAR images using adaptive morphological filter," in *2010 10th International Conference on Intelligent Systems Design and Applications*. IEEE, 2010, pp. 260–265.

[87] B. Zhang, W. Perrie, X. Li, and W. G. Pichel, "Mapping sea surface oil slicks using RADARSAT-2 quad-polarization SAR image," *Geophysical Research Letters*, vol. 38, no. 10, 2011.

[88] J. Lafferty, A. McCallum, and F. C. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," 2001.

[89] T. Liu, X. Huang, and J. Ma, "Conditional random fields for image labeling," *Mathematical Problems in Engineering*, vol. 2016, 2016.

[90] M. J. Collins, M. Denbina, B. Minchew, C. E. Jones, and B. Holt, "On the use of simulated airborne compact polarimetric SAR for characterizing oil–water mixing of the deepwater horizon oil spill," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 3, pp. 1062–1077, 2015.

[91] C. Elachi, "Spaceborne imaging radar: geologic and oceanographic applications," *Science*, vol. 209, no. 4461, pp. 1073–1082, 1980.

[92] H. Espedal, "Satellite SAR oil spill detection using wind history information," *International Journal of Remote Sensing*, vol. 20, no. 1, pp. 49–65, 1999.

[93] H. Espedal and O. Johannessen, "Cover: Detection of oil spills near offshore installations using synthetic aperture radar (SAR)," 2000.

[94] A. H. S. Solberg and E. Volden, "Incorporation of prior knowledge in automatic classification of oil spills in ERS SAR images," in *IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing-A Scientific Vision for Sustainable Development*, vol. 1. IEEE, 1997, pp. 157–159.

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

[96] A. Gambardella, M. Migliaccio, and G. De Grandi, "Wavelet polarimetric SAR signature analysis of sea oil spills and look-alike features," in *2007 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2007, pp. 983–986.

[97] V. V. Malinovsky, S. Sandven, A. S. Mironov, and A. E. Korinenko, "Identification of oil spills based on ratio of alternating polarization images from envisat," in *2007 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2007, pp. 1326–1329.

[98] M. Migliaccio, A. Gambardella, and M. Tranfaglia, "SAR polarimetry to observe oil spills," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 2, pp. 506–511, 2007.

[99] M. Migliaccio, F. Nunziata, and A. Gambardella, "On the co-polarized phase difference for oil spill observation," *International Journal of Remote Sensing*, vol. 30, no. 6, pp. 1587–1602, 2009.

[100] B. Minchew, C. E. Jones, and B. Holt, "Polarimetric analysis of backscatter from the deepwater horizon oil spill using L-band synthetic aperture radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 10, pp. 3812–3830, 2012.

[101] P. C. Genovez, C. E. Jones, S. J. Sant'Anna, and C. C. Freitas, "Oil slick characterization using a statistical region-based classifier applied to UAVSAR data," *Journal of Marine Science and Engineering*, vol. 7, no. 2, p. 36, 2019.

[102] M. M. Espeseth, C. E. Jones, B. Holt, C. Brekke, and S. Skrunes, "Oil-spill-response-oriented information products derived from a rapid-repeat time series of SAR images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 3448–3461, 2020.

[103] S. Skrunes, C. Brekke, and T. Eltoft, "Characterization of marine surface slicks by RADARSATt-2 multipolarization features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5302–5319, 2013.

[104] L. Xu, M. Javad Shafiee, A. Wong, F. Li, L. Wang, and D. Clausi, "Oil spill candidate detection from SAR imagery using a thresholding-guided stochastic fully-connected conditional random field model," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2015, pp. 79–86.

[105] D. I. Morales, M. Moctezuma, and F. Parmiggiani, "Detection of oil slicks in SAR images using hierarchical MRF," in *IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium*, vol. 3. IEEE, 2008, pp. III–1390.

[106] S. Pelizzari and J. Bioucas-Dias, "Oil spill segmentation of SAR images via graph cuts," in *2007 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2007, pp. 1318–1321.

[107] S. Martinis, "Automatic oil spill detection in TerraSAR-X data using multi-contextual Markov modeling on irregular graphs," in *2012 3rd International Conference on Image Processing Theory, Tools and Applications (IPTA)*. IEEE, 2012, pp. 43–46.

[108] F. Parmiggiani, L. Alvarez-Hernandez, and M. Moctezuma-Flores, "A stochastic model for oil spill detection in marine environment with SAR data," in *Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2018*, vol. 10784. SPIE, 2018, pp. 111–116.

[109] H. Guo, D. Wu, and J. An, "Discrimination of oil slicks and lookalikes in polarimetric SAR images using CNN," *Sensors*, vol. 17, no. 8, p. 1837, 2017.

[110] X. Yaohua and M. Xudong, "A SAR oil spill image recognition method based on densenet convolutional neural network," in *2019 International Conference on Robots & Intelligent System (ICRIS)*. IEEE, 2019, pp. 78–81.

[111] B. Minchew, "Determining the mixing of oil and sea water using polarimetric synthetic aperture radar," *Geophysical Research Letters*, vol. 39, no. 16, 2012.

[112] V. Chaudhary and S. Kumar, "Marine oil slicks detection using spaceborne and airborne SAR data," *Advances in Space Research*, vol. 66, no. 4, pp. 854–872, 2020.

[113] J. J. Van Zyl, M. Arii, and Y. Kim, "Model-based decomposition of polarimetric SAR covariance matrices constrained for nonnegative eigenvalues," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 9, pp. 3452–3459, 2011.

[114] S. R. Cloude, D. G. Goodenough, and H. Chen, "Compact decomposition theory," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 1, pp. 28–32, 2011.

[115] V. Chaudhary and S. Kumar, "Dark spot detection for characterization of marine surface slicks using UAVSAR quad-pol data," *Scientific reports*, vol. 11, no. 1, pp. 1–24, 2021.

[116] R. K. Raney, J. T. Cahill, G. W. Patterson, and D. B. J. Bussey, "The m-chi decomposition of hybrid dual-polarimetric radar data with application to lunar craters," *Journal of Geophysical Research: Planets*, vol. 117, no. E12, 2012.

[117] M. Moctezuma and F. F. Parmiggiani, "Adaptive stochastic minimization for measuring marine oil spill extent in synthetic aperture radar images," *Journal of Applied Remote Sensing*, vol. 8, no. 1, p. 083553, 2014.

[118] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Transactions on pattern analysis and machine intelligence*, no. 1, pp. 39–55, 1987.

[119] A. Cheng, M. Arkett, T. Zagon, R. De Abreu, D. Mueller, P. Vachon, and J. Wolfe, "Oil detection in RADARSAT-2 quad-polarization imagery: Implications for scanSAR performance," in *SAR Image Analysis, Modeling, and Techniques XI*, vol. 8179. SPIE, 2011, pp. 119–133.

[120] Z. Zivkovic, "Gentle ICM energy minimization for Markov random fields with smoothness-based priors," *Journal of Real-Time Image Processing*, vol. 11, no. 1, pp. 235–246, 2016.

[121] Y. Boykov and O. Veksler, "Graph cuts in vision and graphics: Theories and applications," in *Handbook of mathematical models in computer vision*. Springer, 2006, pp. 79–96.

[122] D. N. Thomas, *Sea ice*. John Wiley & Sons, 2017.

[123] R. K. Raney, B. Brisco, M. Dabboor, and M. Mahdianpari, "RADARSAT Constellation Mission's operational polarimetric modes: A user-driven RADAR architecture," *Canadian Journal of Remote Sensing*, vol. 47, no. 1, pp. 1–16, 2021.

[124] 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*, vol. 59, no. 12, pp. 9887–9901, 2021.

[125] C. Liu, W. Wang, Q. Zhao, X. Shen, and M. Konan, "A new feature selection method based on a validity index of feature subset," *Pattern Recognition Letters*, vol. 92, pp. 1–8, 2017.

[126] H. Lyu, W. Huang, and M. Mahdianpari, "Eastern Arctic sea ice sensing: First results from the RADARSAT Constellation Mission data," *Remote Sensing*, vol. 14, no. 5, p. 1165, 2022.

[127] Y. Xu and K. A. Scott, "Sea ice and open water classification of SAR imagery using CNN-based transfer learning," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2017, pp. 3262–3265.

[128] 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.

[129] 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 Sensing*, vol. 13, no. 9, p. 1734, 2021.

[130] 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 Sensing*, vol. 13, no. 4, p. 592, 2021.

[131] 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.

[132] Y. Gao, F. Gao, J. Dong, and S. Wang, "Transferred deep learning for sea ice change detection from synthetic-aperture radar images," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 10, pp. 1655–1659, 2019.

[133] C. Bentes, A. Frost, D. Velotto, and B. Tings, "Ship-iceberg discrimination with convolutional neural networks in high resolution SAR images," in *Proceedings of EUSAR 2016: 11th European conference on synthetic aperture radar*. VDE, 2016, pp. 1–4.

[134] 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*, 2016, pp. 770–778.

[135] S. Guo, Y. Tian, Y. Li, S. Chen, and W. Hong, "Unsupervised classification based on h/alpha decomposition and Wishart classifier for compact polarimetric SAR," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2015, pp. 1614–1617.

[136] R. Khedama and A. Belhadj-Aissaa, "Contextual classification of remotely sensed data using map approach and MRF," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 35, no. 7, pp. 11–16, 2004.

[137] 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.

[138] M. M. Espeseth, C. Brekke, and A. M. Johansson, "Assessment of RISAT-1 and RADARSAT-2 for sea ice observations from a hybrid-polarity perspective," *Remote Sensing*, vol. 9, no. 11, p. 1088, 2017.

[139] T. Ainsworth, J. Kelly, and J.-S. Lee, "Classification comparisons between dual-pol, compact polarimetric and quad-pol SAR imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 64, no. 5, pp. 464–471, 2009.

[140] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[141] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[142] 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.

[143] A. K. Qin and D. A. Clausi, "Multivariate image segmentation using semantic region growing with adaptive edge penalty," *IEEE transactions on image processing*, vol. 19, no. 8, pp. 2157–2170, 2010.

[144] R. K. Raney, "Synthetic aperture radar hybrid-polarity method and architecture for obtaining the stokes parameters of a backscattered field," Jun. 29 2010, uS Patent 7,746,267.

[145] 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, 2011.

[146] C. Cortes and V. Vapnik, "Support-vector networks machine learning (pp. 237–297), vol. 20," 1995.

[147] P. C. Dubois-Fernandez, J.-C. Souyris, S. Angelliaume, and F. Garestier, "The compact polarimetry alternative for spaceborne SAR at low frequency," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 10, pp. 3208–3222, 2008.

[148] S. R. Cloude, D. G. Goodenough, and H. Chen, "Compact decomposition theory for L-band satellite radar applications," in *2012 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2012, pp. 5097–5100.

[149] R. K. Raney, "Dual-polarized SAR and stokes parameters," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 3, pp. 317–319, 2006.

[150] A. Bhattacharya, Y. S. Rao, and D. Mandal, *Radar Remote Sensing for Crop Biophysical Parameter Estimation*. Springer, 2021.

[151] N. R. Goodman, "Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction)," *The Annals of mathematical statistics*, vol. 34, no. 1, pp. 152–177, 1963.

[152] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[153] A. Bahri, S. G. Majelan, S. Mohammadi, M. Noori, and K. Mohammadi, "Remote sensing image classification via improved cross-entropy loss and transfer learning strategy based on deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 6, pp. 1087–1091, 2019.

[154] I. Goodfellow, Y. Bengio, and A. Courville, "Deep learning (adaptive computation and machine learning series)," *Cambridge Massachusetts*, pp. 321–359, 2017.

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

[156] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[157] 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 Sensing*, vol. 12, no. 9, p. 1425, 2020.

[158] D. Clausi, A. Qin, M. Chowdhury, P. Yu, and P. Maillard, "Magic: Map-guided ice classification system," *Canadian Journal of Remote Sensing*, vol. 36, no. sup1, pp. S13–S25, 2010.

[159] 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.

[160] F. Mohammadimanesh, B. Salehi, M. Mahdianpari, B. Brisco, and E. Gill, "Full and simulated compact polarimetry SAR responses to canadian wetlands: Separability analysis and classification," *Remote Sensing*, vol. 11, no. 5, p. 516, 2019.

[161] 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.

[162] X. Ma, A. Fu, J. Wang, H. Wang, and B. Yin, "Hyperspectral image classification based on deep deconvolution network with skip architecture," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4781–4791, 2018.

[163] B. Ghimire, J. Rogan, and J. Miller, "Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the getis statistic," *Remote Sensing Letters*, vol. 1, no. 1, pp. 45–54, 2010.

[164] Z. Peng, Z. Guo, W. Huang, Y. Wang, L. Xie, J. Jiao, Q. Tian, and Q. Ye, "Conformer: Local features coupling global representations for recognition and detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

[165] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[166] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[167] J. Yang, C. Li, P. Zhang, X. Dai, B. Xiao, L. Yuan, and J. Gao, "Focal self-attention for local-global interactions in vision transformers," *arXiv preprint arXiv:2107.00641*, 2021.

[168] W. Wang, E. Xie, X. Li, D.-P. Fan, K. Song, D. Liang, T. Lu, P. Luo, and L. Shao, "Pyramid vision transformer: A versatile backbone for dense prediction without convolutions," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 568–578.

[169] X. Chu, Z. Tian, Y. Wang, B. Zhang, H. Ren, X. Wei, H. Xia, and C. Shen, "Twins: Revisiting the design of spatial attention in vision transformers," *Advances in Neural Information Processing Systems*, vol. 34, pp. 9355–9366, 2021.

[170] W. Li, X. Wang, X. Xia, J. Wu, X. Xiao, M. Zheng, and S. Wen, "Sepvit: Separable vision transformer," *arXiv preprint arXiv:2203.15380*, 2022.

[171] P. Zhang, X. Dai, J. Yang, B. Xiao, L. Yuan, L. Zhang, and J. Gao, "Multi-scale vision longformer: A new vision transformer for high-resolution image encoding," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 2998–3008.

[172] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2117–2125.

[173] H. Lin, X. Cheng, X. Wu, and D. Shen, "CAT: Cross attention in vision transformer," in *2022 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2022, pp. 1–6.

[174] L. D. Robertson, H. McNairn, C. McNairn, S. Ihuoma, and X. Jiao, "Compact polarimetry for operational crop inventory," in *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2022, pp. 4423–4426.

[175] H. Wang, C. Xing, J. Yin, and J. Yang, "Land cover classification for polarimetric SAR images based on vision transformer," *Remote Sensing*, vol. 14, no. 18, p. 4656, 2022.

[176] Y. Zhou, H. Wang, F. Xu, and Y.-Q. Jin, "Polarimetric SAR image classification using deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1935–1939, 2016.

[177] Z. Zhang, H. Wang, F. Xu, and Y.-Q. Jin, "Complex-valued convolutional neural network and its application in polarimetric SAR image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 12, pp. 7177–7188, 2017.

[178] H. Dong, L. Zhang, and B. Zou, "PolSAR image classification with lightweight 3D convolutional networks," *Remote Sensing*, vol. 12, no. 3, p. 396, 2020.

[179] S.-W. Chen and C.-S. Tao, "PolSAR image classification using polarimetric-feature-driven deep convolutional neural network," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 4, pp. 627–631, 2018.

[180] C. Yang, B. Hou, B. Ren, Y. Hu, and L. Jiao, "CNN-based polarimetric decomposition feature selection for PolSAR image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 8796–8812, 2019.

[181] W. Xie, G. Ma, F. Zhao, H. Liu, and L. Zhang, "PolSAR image classification via a novel semi-supervised recurrent complex-valued convolution neural network," *Neurocomputing*, vol. 388, pp. 255–268, 2020.

[182] C. Zhang, X. Pan, H. Li, A. Gardiner, I. Sargent, J. Hare, and P. M. Atkinson, "A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 140, pp. 133–144, 2018.

[183] H. Dong, L. Zhang, and B. Zou, "Exploring vision transformers for polarimetric SAR image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2021.

[184] W. Wu, H. Li, X. Li, H. Guo, and L. Zhang, "PolSAR image semantic segmentation based on deep transfer learning—realizing smooth classification with small training sets," *IEEE Geoscience and remote sensing letters*, vol. 16, no. 6, pp. 977–981, 2019.

[185] C. Henry, S. M. Azimi, and N. Merkle, "Road segmentation in SAR satellite images with deep fully convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 12, pp. 1867–1871, 2018.

[186] A. G. Mullissa, C. Persello, and V. Tolpekin, "Fully convolutional networks for multitemporal SAR image classification," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2018, pp. 6635–6638.

[187] A. G. Mullissa, C. Persello, and A. Stein, "PolSARNet: A deep fully convolutional network for polarimetric SAR image classification," *IEEE Journal of selected topics in applied earth observations and remote sensing*, vol. 12, no. 12, pp. 5300–5309, 2019.

[188] Y. Li, Y. Chen, G. Liu, and L. Jiao, "A novel deep fully convolutional network for PolSAR image classification," *Remote Sensing*, vol. 10, no. 12, p. 1984, 2018.

[189] A. Jamali, S. K. Roy, A. Bhattacharya, and P. Ghamisi, "Local window attention transformer for polarimetric SAR image classification," *IEEE Geoscience and Remote Sensing Letters*, 2023.

[190] X. Liu, Y. Wu, W. Liang, Y. Cao, and M. Li, "High resolution SAR image classification using global-local network structure based on vision transformer and CNN," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[191] J. Cai, Y. Zhang, J. Guo, X. Zhao, J. Lv, and Y. Hu, "ST-PN: a spatial transformed prototypical network for few-shot SAR image classification," *Remote Sensing*, vol. 14, no. 9, p. 2019, 2022.

[192] C. Wang, Y. Huang, X. Liu, J. Pei, Y. Zhang, and J. Yang, "Global in local: A convolutional transformer for SAR ATR FSL," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[193] Z.-A. Yang, N.-R. Zheng, and F. Wang, "SAR image classification by combining transformer and convolutional neural networks," in *Proceedings of the 8th China High Resolution Earth Observation Conference (CHREOC 2022) High Resolution Earth Observation: Wide Horizon, High Accuracy*. Springer, 2022, pp. 193–200.

[194] A. B. Ramathilagam, S. Natarajan, and A. Kumar, "TransCropNet: a multichannel transformer with feature-level fusion for crop classification in agricultural smallholdings using Sentinel images," *Journal of Applied Remote Sensing*, vol. 17, no. 2, p. 024501, 2023.

[195] X. Li, L. Lei, Y. Sun, M. Li, and G. Kuang, "Collaborative attention-based heterogeneous gated fusion network for land cover classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 5, pp. 3829–3845, 2020.

[196] Y. Wang, C. M. Albrecht, and X. X. Zhu, "Self-supervised vision transformers for joint SAR-optical representation learning," in *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2022, pp. 139–142.

[197] K. Li, W. Zhao, R. Peng, and T. Ye, "Multi-branch self-learning Vision Transformer (MSViT) for crop type mapping with optical-SAR time-series," *Computers and Electronics in Agriculture*, vol. 203, p. 107497, 2022.

[198] H. Wang, X. Chen, T. Zhang, Z. Xu, and J. Li, "CCTNet: Coupled CNN and transformer network for crop segmentation of remote sensing images," *Remote Sensing*, vol. 14, no. 9, p. 1956, 2022.

[199] Y. Zhang, H. Liu, and Q. Hu, "Transfuse: Fusing transformers and CNNs for medical image segmentation," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24*. Springer, 2021, pp. 14–24.

[200] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101*, 2017.

[201] M. Manore, R. DeAbreu, V. Zabeline, M. Arkett, and D. Bradley, "The radarsat constellation mission (rcm): Extending operational marine surveillance for environment canada," in *Proc. ASTRO Annu. Meeting*, 2010, pp. 1–6.

[202] E. C. Barrett and L. F. Curtis, *Introduction to environmental remote sensing*. Psychology Press, 1999.

[203] J. Li, Y. Pei, S. Zhao, R. Xiao, X. Sang, and C. Zhang, "A review of remote sensing for environmental monitoring in china," *Remote Sensing*, vol. 12, no. 7, p. 1130, 2020.

[204] M. Dabboor, S. Iris, and V. Singhroy, "TheRADARSAT constellation mission in support of environmental applications," in *Proceedings*, vol. 2, no. 7. MDPI, 2018, p. 323.

[205] Q. Yu and D. A. Clausi, "Filament preserving model (FPM) segmentation applied to SAR sea-ice imagery," *IEEE transactions on geoscience and remote sensing*, vol. 44, no. 12, pp. 3687–3694, 2006.

[206] "Interpreting ice charts," *Canadian Ice Service publications, ice forecasts and observations, Environment and Natural Resources, Government of Canada* *https://www.canada.ca/en/environment-climate-change/services/ice-forecasts-observations/publications/interpreting-charts.html*.

[207] C. He, X. Liu, D. Feng, B. Shi, B. Luo, and M. Liao, "Hierarchical terrain classification based on multilayer bayesian network and conditional random field," *Remote Sensing*, vol. 9, no. 1, p. 96, 2017.

[208] F. Qin, J. Guo, and W. Sun, "Object-oriented ensemble classification for polarimetric SAR imagery using restricted boltzmann machines," *Remote sensing letters*, vol. 8, no. 3, pp. 204–213, 2017.

[209] M. Volpi and D. Tuia, "Dense semantic labeling of subdecimeter resolution images with convolutional neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 881–893, 2016.

[210] M. Qi, L. Liu, S. Zhuang, Y. Liu, K. Li, Y. Yang, and X. Li, "Ftc-net: Fusion of transformer and CNN features for infrared small target detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 8613–8623, 2022.

[211] C. Brekke, C. E. Jones, S. Skrunes, B. Holt, M. Espeseth, and T. Eltoft, "Cross-correlation between polarization channels in SAR imagery over oceanographic features," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 7, pp. 997–1001, 2016.

[212] G. Gao, S. Gao, J. He, and G. Li, "Ship detection using compact polarimetric SAR based on the notch filter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 9, pp. 5380–5393, 2018.

[213] S. Sun and C. Hu, "The challenges of interpreting oil–water spatial and spectral contrasts for the estimation of oil thickness: Examples from satellite and airborne measurements of the deepwater horizon oil spill," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 5, pp. 2643–2658, 2018.

[214] F. Charbonneau, M. Arkett, B. Brisco, J. Buckley, H. Chen, D. Goodenough, C. Liu, H. McNarin, J. Poitevin, J. Shang *et al.*, "Meeting canadian user needs with the RADARSAT Constellation Mission's compact polarimetry mode: A summary assessment," *Ottawa. Geomatics Canada Open File*, no. 34, 2017.

[215] U. Kandaswamy, D. A. Adjeroh, and M.-C. Lee, "Efficient texture analysis of SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 9, pp. 2075–2083, 2005.

[216] M. Ohki and M. Shimada, "Large-area land use and land cover classification with quad, compact, and dual polarization SAR data by PALSAR-2," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 9, pp. 5550–5557, 2018.

[217] J. Geng, J. Fan, H. Wang, X. Ma, B. Li, and F. Chen, "High-resolution SAR image classification via deep convolutional autoencoders," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 11, pp. 2351–2355, 2015.

[218] J. Geng, H. Wang, J. Fan, and X. Ma, "Deep supervised and contractive neural network for SAR image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 4, pp. 2442–2459, 2017.