Semiautomated Segmentation of Sentinel-1 SAR Imagery for Mapping Sea Ice in Labrador Coast

Weikai Tan, Jonathan Li, *Senior Member, IEEE*, Linlin Xu, *Member, IEEE*, and Michael A. Chapman

Abstract—This study aims at proposing a semiautomated sea ice segmentation workflow utilizing Sentinel-1 synthetic aperture radar imagery. The workflow consists of two main steps. First, preferable features in sea ice interpretation were determined with a random forest feature selection method. Second, an unsupervised graph-cut image segmentation was performed. The workflow was tested on 13 Sentinel-1A images from January to June 2016, and the results were evaluated on open water segmentation per ice charts provided by Canada Ice Service. The results showed that the proposed workflow was able to segment Sentinel-1 images in to appropriate number of classes, and the potential water identification rate reached 95%.

Index Terms—Image segmentation, sea ice, sentinel-1, synthetic aperture radar.

I. INTRODUCTION

ARCTIC sea ice extent has been decreasing rapidly through recent decades [1], [2] with a recorded loss of over 1 million km^2 compared to the historical average from late 20th to early 21st century [3]. In addition to loss of coverage, some evidence of ice thickness reduction has also been observed [4], [5]. If the decreasing trend continues, it is possible that perennial ice in the Arctic might vanish within a few decades per several climate models from Intergovernmental Panel on Climate Change 4th Assessment Report (IPCC AR4) [6], [7]. The continuous loss of sea ice is considered to have profound impacts on Arctic climate, hydrological cycle, and ecology locally, regionally and globally [8]–[11]. In addition to the natural environment, human activities are greatly affected by Arctic sea ice dynamics as well. Marine activities in ice-infested regions such as the Canadian northern and eastern coast are highly dependent on navigable routes. According to Canadian Coast Guard, the number of ships navigated through Canadian Arctic has tripled the number since 1990 [12]. The shrinking of Arctic sea ice may provide longer ice-free periods and more navigable routes, resulting in potential marine traffic increase in the future.

Given the significant impact of sea ice on climate and human activities, it is crucial to monitor sea ice extents and conditions in a timely and accurate manner. Satellite synthetic aperture radar (SAR) systems are suitable for sea ice mapping since they are active sensors that can acquire data regardless of solar illumination and atmospheric conditions, enabling reliable and long-term monitoring of sea ice. In addition, polar regions usually have long cloudy periods, which may result in temporal data acquisition gaps for optical sensors [13]. More frequent cloud coverage in Arctic regions can be expected in the near future, as precipitation has been observed to significantly increase in recent decades [14]. Canadian Ice Service (CIS), the official provider of sea ice information in Canada, has been using SAR imagery in daily operations for almost 30 years, but there are two major challenges in sea ice monitoring.

First, CIS has been heavily relying on RADARSAT-2 since RADARSAT-1 stopped service in 2013, and other sources thus need to compensate for data limitation or potential loss of RADARSAT-2. Sentinel-1 is a SAR satellite mission consists of two satellites equipped with SAR sensors, and the images are open to public. Terrain Observation Progressive Scan (TOPS) acquisition mode [15] with expected better performance compared to conventional ScanSAR mode [16] is used in large-swath mapping. The first challenge is to find out whether Sentinel-1 can provide satisfying performance in ice products generation. Second, CIS received about 64 000 RADARSAT-1 images from 2006 to 2013 and approximately 43 000 RADARSAT-2 images from 2007 to 2014 [17]. The interpretation of images and sea ice product generation still rely on manual processing of experts, and algorithms that process SAR images automatically or with less human supervision are desirable. The main goal of this study is to explore the effectiveness of sea ice monitoring using Sentinel-1 imagery with a proposed semiautomated image segmentation workflow to overcome the abovementioned two challenges.

Various studies have been published recently on sea ice monitoring systems using satellite SAR sensors. Pixel-based image classifications are effective in sea ice image interpretation [18], [19], but region-based methods involving spatial features have shown the advantage in suppressing speckle noise and

W. Tan is with the Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: weikai.tan@uwaterloo.ca).

J. Li is with Fujian Key Lab of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, Xiamen 361005, China and also with the Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: junli@xmu.edu.cn).

L. Xu is with the School of Land Science and Technology, China University of Geosciences, Beijing 430074, China, and also with the Department of Systems Design Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: l44xu@uwaterloo.ca).

M. A. Chapman is with the Department of Civil Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: mchapman@ryerson.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2018.2806640

salt-and-pepper effect. One of the successful systems is the Map-Guided Ice Classification (MAGIC) software [20], which is currently being considered by CIS in operational sea ice monitoring [17]. The MAGIC software first segments SAR images with initial labels, and an iterative region-growing using semantics (IRGS) method [21] is applied to merge adjacent regions with the same labels. MAGIC has demonstrated its potential of operational sea ice mapping with single and dual polarized RADARSAT-2 imagery [22]–[25]. In addition to IRGS, other segmentation methods such as enhanced total variation optimization segmentation [26] and iterated conditional modes (ICM) [27] have also been demonstrated to be effective. There is also advanced weakly supervised classification algorithm that is built upon IRGS for dealing with cases where the number of training samples is limited [28]. Moreover, there are other methods utilizing full polarization [29] or deep learning techniques [30], but coverage limit of full-polarized SAR data or demand of large training sets made these methods difficult to implement. Therefore, region-based segmentation could be more effective and practical for processing large-sized sea ice SAR images. A typical region-based segmentation workflow consists of three steps: 1) image preprocessing, 2) feature extraction, and 3) image segmentation.

One of the main challenges in interpreting SAR images is inevitable speckle noise. The pixel-based methods usually apply speckle filters to suppress noise during preprocessing [29], [31], whereas region-based methods are designed to reduce the influence of speckle noise [26]. The operational sea ice monitoring by CIS does not require processing at full resolution [32]; therefore, reducing image size is reasonable and efficient, and speckle noise can be suppressed by pixel averaging [18], [31] and more advanced denoising approach [33]. In addition to speckle noise, backscatter variation across wide-swath images is another challenge, as incidence angle difference between near and far end of each image is not negligible. For example, in RADARSAT-2 dual-polarized imagery, backscatter intensity decreases as incidence angle increases in HH bands, while HV bands have "banding" effect between swaths. Linear correction could be performed in copolarized bands during the preprocessing stage if certain wind and wave conditions could be met [13], [18], [34]. Banding effects in cross-polarized bands could be reduced by separately processing different beams [19], [31]. However, the differences vary on different surface, and ground truth is not always available especially before image classification. To overcome statistical nonstationarities of within a large scene, region-based methods seems to have advantages even if no correction is performed [24]. However, backscatter variance in Sentinel-1 images has different characteristics, the performance of previous methods based on ScanSAR imagery is not guaranteed.

Spatial features are considered to be effective in sea ice studies, and gray-level cooccurrence matrix (GLCM) features are the most popular ones. Various features and window sizes have been tested in [18], [19], and [31], but no consensus has been made. Feature selection methods such as forward searching [24] or dimension reduction methods such as principal component analysis (PCA) [25], [35] can be performed to produce most effective

features. Though PCA is able to reduce dimensions of features, all the features still need to be calculated prior to dimension reduction, which may not be efficient for automated systems. As a result, a forward searching feature selection method using random forests [36] is adopted in this study.

Previous studies have shown advantages of image segments over pixels. IRGS, the image segmentation method used in MAGIC, can be considered the most acknowledged technique in sea ice studies, but one shortcoming is that the number of classes is predetermined. The determination of the number could be tricky, as the suitable number of classes in different images are not the same. Similar to IRGS, optimization can also be achieved by graph cut [37], and this method has been extended by adding label cost to reduce number of labels automatically [38]. Limited applications of this method have been found in remote sensing studies [39]–[41], especially for SAR imagery. Graph cut is demonstrated to be more effective in image segmentation compared with traditional methods such as ICM [42], and found to have better performance in some case studies compared with K-means, Iterative Self-Organizing Data Analysis Technique [39], Gaussian mixture model and region growing [40]. Therefore, we believe that this label optimization method with the advantage of determining number of labels automatically could help us be one more step approaching automated sea ice classification.

II. Study Area and Data

The study site of this study is Labrador coast, the mainland part of the Canadian province Newfoundland and Labrador, which approximately locates between 51.9° and 60.6° N, 55.4° and 64.6°W. Labrador Sea and Baffin Bay covers approximately 20% of ice coverage in the northern hemisphere, and it has been found to be one of the regions where sea ice coverage decreases most during the past 30 years [43]. Sea ice variability in Labrador Sea has direct impacts on climate [44], as well as marine transportation for Canada [45]. The ice regime in Labrador coast is heavily affected by winds and currents, especially by the famous Labrador Current, making it difficult to interpret. Normally only first-year ice is present during winter time, while on few occasions will old ice occur [46]. Ice usually freezes up before mid-December, and it usually clears out before August. The ice types are mainly new ice and first-year ice, but the spread of ice from shore varies year to year depending on the low-pressure system present in winter. In this study, sea ice in the marginal ice zone (MIZ) of Labrador coast, which has one of the most difficult ice regimes, during the winter from late-2015 to mid-2016 was investigated.

As listed in Table I, a total of 15 scenes from Sentinel-1A in EW mode were used in sea ice segmentation. The EW mode has the largest scene size among all beam modes of Sentinel-1, with a swath width of 400 km. The Level-1 Ground Range Detected Medium (GRDM) product was used, which means the products have been focused, multilooked, and georeferenced into the World Geodetic System 1984 (WGS84). The pixel spacing and the scene size are approximately 40 m and $10\,000 \times 10\,000$ pixels, respectively. All scenes are in dual

TABLE I
DATASET

| Dataset ID	Date	Acquisition time	Direction
1	2015-12-28	10:27	Descending
2*	2016-01-04	10:20	Descending
3	2016-01-11	10:12	Descending
4*	2016-01-18	10:03	Descending
5	2016-01-21	10:28	Descending
6	2016-02-04	10:12	Descending
7	2016-02-14	10:27	Descending
8*	2016-03-16	10:20	Descending
9	2016-03-23	10:11	Descending
10*	2016-04-02	21:41	Ascending
11	2016-04-09	21:33	Ascending
12	2016-04-26	21:40	Ascending
13	2016-05-20	21:40	Ascending
14	2016-05-27	21:34	Ascending
15	2016-06-13	21:42	Ascending

*Datasets used for training
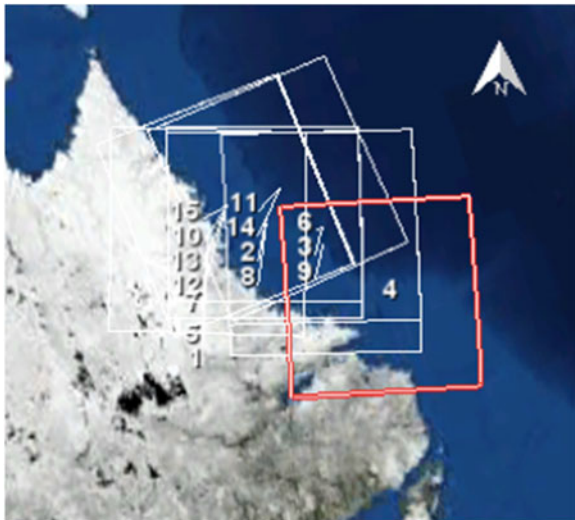
Fig. 1.   Image footprints of 15 Sentinel-1 scenes. (Red box showing image on January 18, 2016).

polarization mode containing HH and HV. The incidence angle ranges from 18.9° to 47.0°. Fig. 1 shows the image footprints of all 15 scenes, covering Labrador coast, where most coastal human activities happen. The highlighted footprint is the scene from January 18, 2016, which is the scene used for illustration purpose in later sections. In addition to Sentinel-1 imagery, ice charts provided by CIS were used as ground truth.

III. METHODS

A. Image Preprocessing

Some preliminary processing has been done during the production of the Sentinel-1 GRDM product, some further processes are needed in this proposed workflow. The preprocessing was done using Sentinel Application Platform software produced by the European Space Agency (ESA). Fig. 2 shows the modules used in the preprocessing step.
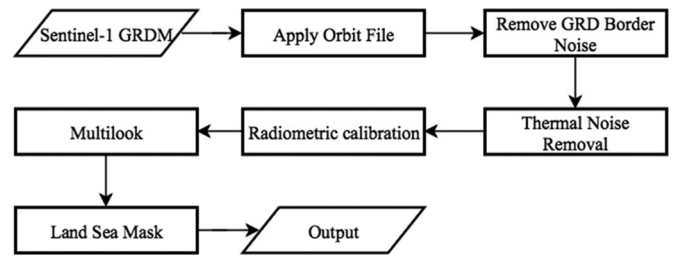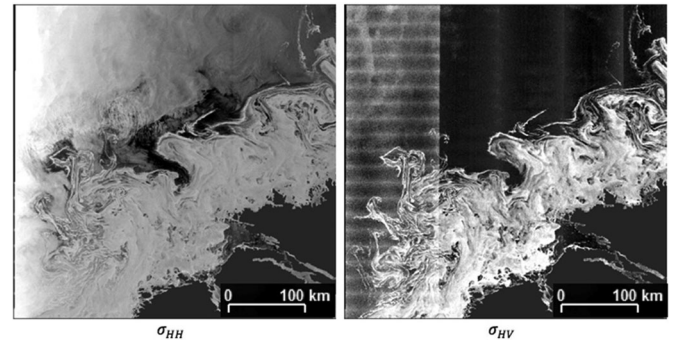

Fig. 2.   Flowchart of the proposed method.

Fig. 3.   Sentinel-1 scene on January 18, 2016. (σ_0 values stretched to 0–255).

The first step is retrieving orbit files from the server. Although orbit parameters are already in the GRDM products, more precise orbit parameters can be retrieved from ESA approximately two weeks after the generation of Sentinel-1 products. Thereafter, border noise removal was performed since some artifacts may present during the generation of Sentinel-1 GRDM products. Thermal noise removal was performed according to a look-up-table within the GRDM product, but the process is not ideal on ocean surface for the time being. The radiometric calibration transforms magnitude into sigma nought (σ_0) values, which is a measurement of radar backscatter on the ground. The magnitude values were transformed in decibel scale.

According to [13] and [30], 1000×1000 pixels would be representative in processing full-scene images for prototype algorithms. The multilooking process at this stage worked as pixel averaging by an 8×8 window, resulting in an image size of approximately 1250×1250 pixels. Finally, as the target of interest is sea ice, land regions were masked out using Shuttle Radar Topography Mission Digital Elevation Model (DEM) 3-arc-second product. In addition, the cross-polarization ratio (σ_{HH}/σ_{HV}) is considered effective in sea ice studies in [13], so that the ratio was calculated.

Fig. 3 shows the scene captured on January 18, 2016, which represents ice conditions in Labrador coast during freeze-up time, and only part of the image is covered by ice. It can be easily identified that the middle part in the scene is covered by ice, and the top left half of the scene is covered by water or very thin ice. Since C-band microwave has a wavelength of approximately 5 cm, very thin ice on top of water with thickness less than 5 cm could be easily penetrated by microwave. As a result, the thin ice may not be distinctive from water in Sentinel-1 imagery. In this study, very thin ice is grouped into the class of

"water," and "water" used in the following sections may contain part of very thin ice which might not be identifiable. Visually, HV has better contrast between different classes, while HH may capture some characteristics at the surface. For example, in the middle-left part in the image, some structures of ice surface could be found in HH, while in HV those do not present. Both HH and HV contribute to the SAR image interpretation, and by using only one of them may result in misinterpretation.

The noise equivalent sigma zero (NESZ) of Sentinel-1 imagery is around -22 dB [47], and values close or below this level could be too noisy to be useful, especially in cross-polarized bands [48]. However, in sea ice studies, water and several ice types may be close to the NESZ or even lower, which introduces a challenge for image interpretation. Water is the most easily identified class in HV, which has very low backscatters, but has strong banding effects, especially in the first subswath. Banding effects also present in RADARSAT-2 and other SAR images, but backscatters are mostly even within each subswath. However, repeating horizontal lines (along look direction) and beam seams in HV [47], especially in the first subswath, presents in Sentinel-1 imagery. The noise came from the background as little or no signal could be received in the areas below NESZ. Though thermal noise removal has been performed, current solution does not provide satisfying results. In these noise-contaminated regions, water and some ice classes could be easily misinterpreted only from raw σ_{HV} values. Moreover, water is affected by incidence angle effects, wave and wind conditions, resulting in backscatter variances in Sentinel-1 imagery.

Preliminary tests showed that noise in the first subswath in cross-polarized bands prevents successful identification of ice types and water when full-scene images were processed. Some attempts have been made to reduce the noise [49], [50], but the methods were still not ideal before ESA produces an official solution [51]. Therefore, the first subswath was removed as a compromise to reduce influence of the noise pattern in Sentinel-1 imagery. Such noise also presents in the remaining part of the images, but the variations were observed to be smaller and more even across the other swaths. As a result, the image size was further reduced to approximately 915×1250 pixels.

B. Training Sample Selection

To determine the effective features in identifying ice types and water, training samples are selected for evaluation. Four scenes were chosen to select training samples for feature selection as they contain most number of ice classes, and the corresponding acquisition dates were January 4, January 18, March 16, and April 2, 2016. Before January, ice just started to form so that few ice classes present, and after April ice started to melt and fewer classes could be identified. The selection of ice classes was based on ice charts provided by CIS, as well as visual interpretation.

Fig. 4 shows an RGB composite of the scene on January 18, 2016 and the selected training samples. The scene was not projected so that left and right are reversed compared to the image footprint. By interpretation of the corresponding ice chart,
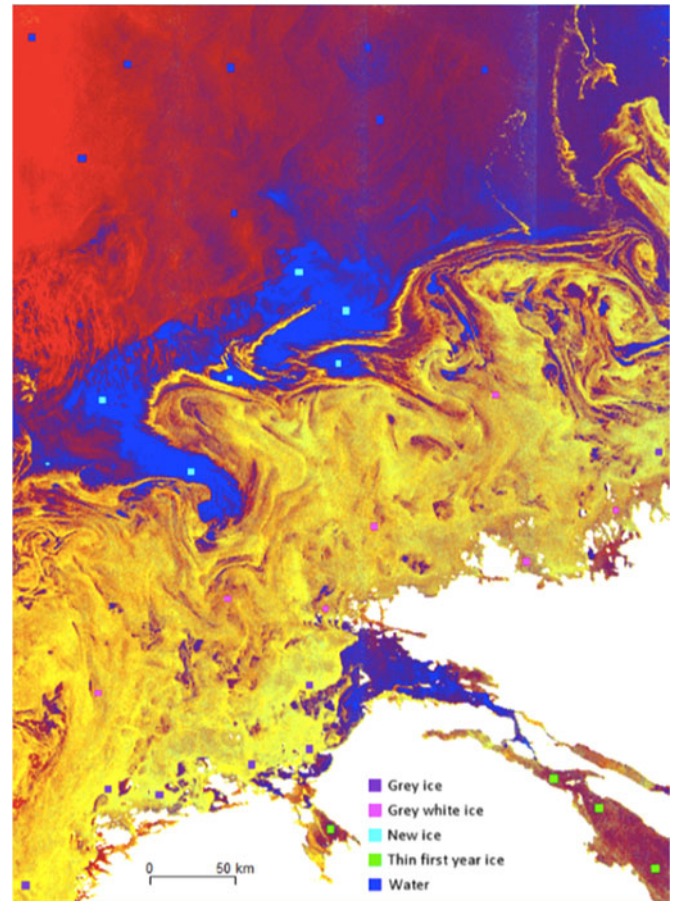

Fig. 4. Sentinel-1 scene in RGB (R: σ_{HH}, G: σ_{HV}, B: σ_{HH}/σ_{HV}) composite on January 18, 2016 and training samples.

most ice in this scene are gray ice and gray white ice, while the lower right corner and top right corner of the image is covered by thin first year ice. In addition, new ice also presents at MIZ. Training samples were selected in small rectangles covering the inferred classes, while complicated regions such as the mid-left part were avoided to assure "purity" of the samples. In addition, the training samples were spread out to cover different incidence angles to capture incidence angle effects. The training samples were selected in ENVI. In this scene, 3163 pixels were selected as training samples. Approximately 3000 pixels were also selected in each of the other three scenes.

C. Feature Selection

Spatial features have demonstrated its effectiveness in interpreting remote sensed images, and as previously summarized, GLCM features [52] has the most popularity. GLCM probabilities represent all pairwise combinations of gray levels in the window of interest, and the textures are determined by three parameters: window size, interpixel distance, and orientation. From the literature on GLCM features [24], multiple window sizes and interpixel distances have been selected to calculate GLCM features on σ_{HH}, σ_{HV}, and σ_{HH}/σ_{HV}. Window size 3, 5, 11 with step size 1, window size 25 with step size 1, 5, and window size 51 with step size 5, 9 were selected. In terms of

orientation, the look direction is considered to perform slightly better in SAR images [53]. Adding that a linear decreasing trend was observed between σ_{HH} and incidence angle in water, the look direction was selected with the expectation of reducing the effect. Eight texture measurements: mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation were calculated, resulting in a total of 168 candidate GLCM features. The texture features were produced in R using package "GLCM" [54]. The texture features as well as the original bands were normalized to 0–255. In addition to the 168 texture features, original σ_{HH}, σ_{HV}, and $\sigma_{\mathrm{HH}}/\sigma_{\mathrm{HV}}$ values were taken into the next step, resulting in 171 features.

A random forest feature selection method utilizing forward searching [36] was adopted to select the most representative feature. The main idea of random forest [55] is to combine a number of decision trees built from bootstrap samples in the training set using a random subset of variables. Random forest is suitable for feature selection as it contains an importance measurement during the process. In addition, random forest could achieve better performance and be more computationally efficient compared to SVM [56]. The feature selection process was done in R with package "VSURF" [57]. The feature selection process contains three steps: 1) thresholding, 2) interpretation, and 3) prediction. Detailed description and theoretical basis could be found in [36].

In the first step, all the n variables were ranked by variable importance in the descending order, and the least important variables were removed with m variables left. In this case, variable importance is embedded in the random forest classification process, and it is determined by out-of-bag (OOB) error. During the process of random forest classification, each tree t is created from a subset of training data, while the data not selected are called OOB samples, and the misclassification rate on the OOB samples using tree t is called OOB error ($E_{\mathrm{oob},t}$). The features are denoted as X and the ith ($i \in [1:171]$) feature is denoted as X^i. To evaluate the contribution of X^i, the values of X^i in sample OOB_t are permuted with random values, and the classification error rate using tree t in this new sample OOB'_t is denoted as $E'^{\ i}_{\mathrm{oob},t}$. The difference between the two error rates evaluates how important X^i is in tree t. The importance of variable X^i is defined as

$$VI\left(X^i\right) = \frac{1}{n_{\mathrm{tree}}} \sum_t \left(E'^{\ i}_{\mathrm{oob},t} - E_{\mathrm{oob},t}\right) \qquad (1)$$

where n_{tree} is the total number of trees built in this random forest, and the variable importance is the average error rate difference. The larger the mean error rate, the variable is more important. All the n variables were ranked by average importance through 50 runs and the least important variables were removed based on a threshold T_1 given based on a Classification and Regression Tree (CART) model [58] on standard deviation of variable importance. Variables with average importance values lower than T_1 were removed, resulting in m important variables.

In the second step, the smallest k variables ($k = 1$ to m) that produced adequately low error rate was selected. It is believed that using all m variables leads to the lowest E_{oob}, and the threshold was set as the lowest mean E_{oob} plus its standard

TABLE II
RESULTS OF FEATURE SELECTION

Most important variables	Rank
σ_{HH} variance 11×11 Step 1	1
σ_{HV} contrast 25×25 Step 5	2
σ_{HV} mean 11×11 Step 1	3
σ_{HV} correlation 25×25 Step 1	4
σ_{HH} variance 25×25 Step 5	5
σ_{HV} dissimilarity 25×25 Step 5	6

deviation over 25 runs. Random forests were built started from $k = 1$ with the most important variable to $k = m$, and when the mean E_{oob} reaches the threshold T_2, the smallest set was determined.

In the third step, based on the importance ranking of the k selected variables, an ascending sequence of variables were used to form random forests where only if mean E_{oob} was significantly reduced. The threshold T_3 is determined by the mean E_{oob} difference with variables left out at the previous step, as denoted by

$$T_3 = \frac{1}{m-k} \sum_{j=k}^{m-1} |E_{\mathrm{oob}}\left(j+1\right) - E_{\mathrm{oob}}(j)| \qquad (2)$$

where $E_{\mathrm{oob}}(j)$ represents the mean E_{oob} using j most important variables. The step started from $l = 1$, which is the most important variable, and only if $E_{\mathrm{oob}}(l+1) - E_{\mathrm{oob}}(l)$ was larger than the threshold, the $l + 1$th variable was kept. Thus, less important variables and variables with high correlation could be eliminated. After the three steps, a minimal set of l variables were determined.

One-hundred GLCM features, as well as original bands (σ_{HH}, σ_{HV}, and $\sigma_{\mathrm{HH}}/\sigma_{\mathrm{HV}}$), with a total of over 12 000 samples, were put into feature selection. Since high correlations may exist between GLCM features with different window sizes, adding that only a small set is needed in the image segmentation step, an aggressive parameter selection could improve computational efficiency. The first parameter affecting computation speed is n_{tree}, the number of trees in random forest. According to the author's previous study [59], 20 would be sufficient for SAR image classification. Another parameter setting that influences efficiency is that three thresholds (T_1, T_2, T_3) can be multiplied by coefficients (C, C_2, C_3) considering different circumstances. With a larger coefficient, fewer variables would be left in each step. The three coefficients were set according to several testing results. Finally, 6 variables, as shown in Table II, were finally selected, and these features were calculated for all 15 scenes to perform image segmentation.

D. Proposed Image Segmentation Workflow

In this study, a semiautomated workflow is proposed to segment full scene Sentinel-1 imagery into reasonable number of classes. There are three main steps: 1) image preprocessing, 2) feature extraction, and these features were determined by feature selection, and 3) image segmentation with desired parameter settings. The first two steps were introduced in the previous

sections. The images were calibrated into σ_0 values, and land regions and the first subswath were removed. Afterward, GLCM features were extracted on the preprocessed images, and image segmentation would be performed on these features.

In sea ice studies, because of the large in-class variation of the ice classes, it is difficult to determine the appropriate number of classes in image segmentation in either supervised or unsupervised methods. For example, water suffers incidence angle effects, adding that the response of water may be different at different wind speeds, so that the variation of backscatter values may be larger than the difference of that between two ice types, such as gray ice and gray white ice. From the perspective of image classification, it might be more appropriate to assign several class labels to one single class with large variations, especially when an unsupervised method is taken. IRGS [24], the state-of-the-art algorithm, assigns six labels in segmentation step before automatically classify one scene into two classes, ice and water. In this study, we intended to propose a workflow that can segment SAR images into a larger number of classes, so that selecting a suitable number of classes is crucial. In addition, during winter time there might be five or more ice classes, while only one or two types of ice appears during freeze-up and melting periods; thus, whether one single selection of number of classes can be applicable in different scenes becomes a challenge. An optimization algorithm utilizing label cost, which is capable of determining the number of labels dynamically, was chosen in this study. The detailed description of this method can be found in [38].

Image segmentation can be transformed into a multilabel optimization problem, and three costs are taken into consideration: 1) data cost, 2) smooth cost, and 3) label cost. The total energy E of a set of label l can be denoted as

$$E(l) = w_d \cdot \overbrace{\sum_{p \in P} D_p(l_p)}^{\text{data cost}} + w_s \cdot \overbrace{\sum_{q \in N_p} V_{pq} \cdot (l_p, l_q)}^{\text{smooth cost}} + w_l$$
$$\cdot \overbrace{\sum_{l \subseteq L} h_l \cdot \delta(l)}^{\text{label cost}} \tag{3}$$

where w_d, w_s, and w_l stands for the weights of data cost, smooth cost, and label cost, respectively. These three weights are relative weights, so that w_d was set to 1. A larger w_s promotes smoothness but boundaries between classes may become unclear, and a larger w_l encourages fewer classes but some classes covering small areas may be lost. w_s is referred as scale, while w_l is referred as label cost in the rest of the paper.

Data cost measures how well the assigned label l_p of a pixel p, which minimizes in-class variation of the variables. The cost was defined as the Euclidian distance to cluster centre

$$D_p(l_p) = \sqrt{\sum (x_p - \mu_l)^2} \tag{4}$$

where x_p represents all values of the six variables at pixel p and μ_l represents mean values of all pixels with label l.

Smooth cost measures label consistency between neighboring pixels, which reduces typical "salt and pepper" problem in image classification. In this study, the eight adjacent pixels around pixel p were defined as neighbors. A "smoother" segmentation result encourages neighboring pixels to have the same label. If pixel q is within the neighborhood N_p of pixel p, the function V_{pq} penalizes the segmentation if $l_p \neq l_q$. Any class could border any other class and no specific priori was defined, so V_{pq} was defined as

$$V_{pq} = \begin{cases} 1, & l_p \neq l_q \\ 0, & l_p = l_q \end{cases} . \tag{5}$$

Label cost penalizes excessive number of labels to promote data compactness. The indication function $\delta(l)$ was defined as

$$\delta(l) = \begin{cases} 1, & \exists p : l_p \in l \\ 0, & \text{else} \end{cases} . \tag{6}$$

The order of label l was not preset, and each label does not have a specific class name, so that all the labels were given the same penalty if exists by setting the coefficient h_l as 1. During the process of image segmentation, one or more labels might be merged with other labels if a subset l was found in the initial label set L that lead to smaller total energy.

The energy minimization problem with data cost, smooth cost, and label cost can be solved using graph cut, a method of partitioning vertices of a graph to achieve energy minimization [37]. In this study, an extended version of the α-expansion method [38] was used to achieve energy optimization. The α-expansion method is able to switch labels of a large number of pixels simultaneously via graph cut, which results in faster approximation to energy minimization compared to traditional pixel-by-pixel swap. The workflow of α-expansion used is as following: 1) start with initial labeling L, 2) α-expansion on each label in L and find the minimum $E(l)$, 3) if $E(l) < E(L)$, $L = l$, save labeling result and iterate 2) and 3). During this process, some labels may be merged into the α label, and the number of labels could be reduced.

The initial labeling was produced by K-means unsupervised labeling, and the optimization process kept iterating until less than 1% of the pixels changed averaged by the three last iterations, or the number of iterations reached 100. Thus, final segmentation map was produced. The segmentation was done using software "GCoptimization" [60] in MATLAB.

E. Accuracy Assessment

Since egg codes of ice charts do not have clear boundaries of each ice type, it is not feasible to evaluate the segmentation accuracy for different ice types. As a compromise, accuracy assessment can only be performed to evaluate the correctnesss of distinguishing ice and water, two most obvious classes, using this proposed segmentation workflow.

There are three classes: water, ice, and land present in the 15 scenes, and land can be determined by DEM or other maps; thus, ground truth for either ice or water is needed. However, after removal of the first subswath, only 13 scenes contain water. The ground truth of water was generated by manual digitization based on the interpretation of SAR images with reference to ice

<div align="center">

TABLE III
ACCURACY ASSESSMENT FOR IMAGE SEGMENTATION

</div>

Date	Label cost 15*					Label cost 20					Label cost 25				
	UA	PA	OA	NT	NW	UA	PA	OA	NT	NW	UA	PA	OA	NT	NW
0104	75.4%	83.6%	94.7%	6	1	47.5%	85.6%	86.8%	5	1	47.5%	85.6%	86.8%	5	1
0111	90.1%	96.6%	96.0%	7	2	90.1%	96.6%	96.1%	7	2	91.8%	96.1%	96.5%	4	1
0118	61.8%	99.8%	85.4%	5	2	61.6%	99.8%	85.3%	5	2	61.9%	99.8%	85.5%	5	2
0121	85.3%	91.2%	97.9%	4	1	29.3%	99.9%	79.5%	3	1	29.3%	99.9%	79.5%	3	1
0204	88.3%	99.4%	96.0%	6	2	87.9%	99.3%	95.9%	5	2	88.0%	99.3%	95.9%	5	2
0316	14.2%	88.2%	86.9%	5	1	14.2%	88.8%	86.7%	5	1	12.4%	89.9%	84.5%	4	1
0323	95.3%	95.9%	97.0%	8	3	96.2%	95.2%	97.0%	7	2	85.9%	98.1%	93.8%	4	2
0402	97.6%	86.9%	95.8%	5	1	93.4%	91.1%	95.8%	4	1	93.4%	91.1%	95.8%	4	1
0409	96.7%	98.7%	97.4%	4	1	96.7%	98.7%	97.4%	4	1	96.7%	98.7%	97.4%	4	1
0426	88.7%	84.2%	95.2%	5	1	87.8%	84.2%	95.1%	4	1	78.3%	86.4%	93.3%	3	1
0520	90.9%	99.6%	95.4%	4	1	86.5%	99.9%	93.1%	3	1	86.5%	99.9%	93.1%	3	1
0527	97.1%	98.6%	97.0%	5	2	93.0%	99.0%	94.1%	4	2	93.0%	99.0%	94.1%	4	2
0613	97.8%	99.5%	98.3%	3	1	97.8%	99.5%	98.3%	3	1	97.8%	99.5%	98.3%	3	1
Overall	**88.4%**	**96.6%**	**94.9%**			**82.4%**	**97.2%**	**92.4%**			**81.2%**	**97.5%**	**91.9%**		

PA: producer's accuracy; UA: user's accuracy; OA: overall accuracy; NT: number of total labels; NW: number of water labels.

*15 was finally selected as label cost.

charts provided by CIS. A ground truth image with two classes: water and others, was produced for 13 scenes.

The results of image segmentation may contain different number of classes ranging from 3 to 8, and this proposed method does not include a classification step due to uncertainties of sea ice types. Therefore, a pixel-by-pixel accuracy assessment was conducted with manual labeling of ice and water. The possible water segments were manually picked and merged, and the others were also merged accordingly. An error matrix [61] was produced for each of the segmentation scenarios, and overall accuracy, user's accuracy, and producer's accuracy of class water were calculated.

IV. RESULTS

There are three main parameters in the segmentation algorithm: 1) initial number of classes K, 2) w_s as the weight of smooth cost, represented by "scale", and 3) w_l as the weight of label cost.

K determines the maximum number of classes during the segmentation process, and tests starting with different values of K were conducted. Generally, the final number of classes is determined by the selection of label cost, and similar results were achieved with different values of K. Thus, K should be set according to a desired final number of classes. In this study, only the separation of ice and water would be assessed, so that the optimal final classes would be as water, ice, and land. To successfully separate ice and water, six classes were used for initial segmentation in IRGS to fit different conditions [24]. Therefore, the final number of classes in this study would target at around 6 or a smaller number. Finally, 10 was chosen for K at initial K-means labeling after testing to assure completeness of class scheme, and to validate the performance of label cost.

The second parameter, scale, determines the smoothness of the segments. With a larger scale, the results would become smooth but some details might not be preserved, while with a small scale, small fragments would remain which affects interpretation. Generally, scale affects the results more from a

visual perspective, and a larger scale would promote the merge of classes during iteration. By visual comparison, scale was set to 20 to achieve a sensible smoothness of segments.

The third parameter, label cost, is the most important parameter in this study, which determines how aggressively number of labels is reduced. Generally, a larger label cost promotes smaller number of labels, but may inappropriately merge different classes if the cost is too large. In addition, depending on the distribution of pixel values in different scenes, the same label cost setting may result in different number of labels during image segmentation.

After testing with different label costs, 15, 20, and 25 were considered to be candidates of parameters. The results of accuracy assessment using the three label cost settings are listed in Table III. Generally, all three label cost settings generated sensible results in distinguishing water from other classes, with the overall accuracy of 94.9%, 92.4%, and 91.9%, with label cost settings as 15, 20, and 25, respectively. By comparing the three label cost settings, 15 generated the highest user's accuracy and overall accuracy. The label cost of 20 led to slightly higher producer's accuracy but lower user's accuracy, while reduced number of labels could be observed. The label cost setting of 25 further reduced the number of labels, and the user's accuracy and producer's accuracy also decreased, but the producer's accuracy slightly increased. The larger label cost settings may result in excessive merging of some classes so that producer's accuracy increases at the cost of user's accuracy.

By comparing the producer's and the user's accuracy, it can be observed that producer's accuracy is higher than user's accuracy in all three cases, which means that certain types of sea ice may have very similar response to water. However, the producer's accuracies could reach over 96% percent, indicating that most of the water in the images could be identified. For demonstration purpose, four of the segmentation results with the highest overall accuracy is shown in Fig. 5.

The workflow was tested on a personal computer with dual-core 1.8 GHz Intel Core i5 processer and 8G memory. The
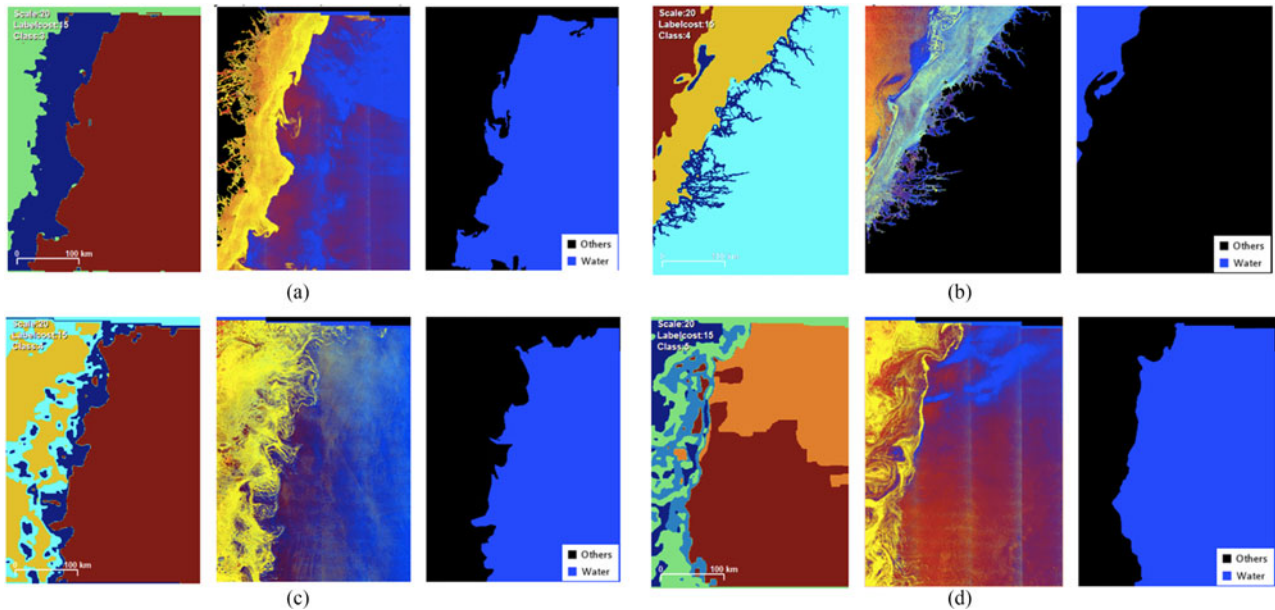
Fig. 5. Sample image segmentation results with scale: 20, label cost: 15. (PA: producer's accuracy; UA: user's accuracy; OA: overall accuracy.) Left: segmentation result, Middle: RGB (R: σ_{HH}, G: σ_{HV}, B: σ_{HH}/σ_{HV}) combination of SAR images, Right: ground truth. (a) Scene: June 13; UA:97.8%; PA: 99.5%; OA: 98.3%. (b) Scene: January 21; UA:85.3%; PA: 91.2%; OA: 97.9%. (c) Scene: April 9; UA:96.7%; PA: 98.7%; OA: 97.4%. (d) Scene: May 27; UA:97.1%; PA: 98.6%; OA: 97.0%.

feature selection algorithm took approximately 2 h with the final coefficient settings, and it could be reduced with a more powerful CPU. Nevertheless, feature selection only needs to be processed once. The segmentation process took approximately 30 min per iteration with testing weight settings, and the processing time was different for each image. Generally, it took less than ten iterations to converge. Since the segmentation algorithm is a single-thread computation, the processing time could be reduced with a more powerful CPU and multiple images could be processed in parallel with multiple cores.

V. DISCUSSION

A. Noise in Sentinel-1 Images

As mentioned previously, backscatters close or below the noise floor show some variations in Sentinel-1 imagery of sea ice and water, and the noise from the sensor may result in misinterpretation of the imagery. The backscatter variance on the surface with lowest values would show some noise patterns.

Fig. 6 illustrates the backscatter variation on water at different incidence angles in the scene on January 18, 2016, where pixels of water were manually selected. σ_{HH} of water has very large variation from around 0 to -30 dB, which may largely affect classification purely from σ_{HH} values. However, an obvious linear trend could be found as incidence angle increase σ_{HH} of water decrease, which corresponds well with previous studies on ScanSAR imagery.

In terms of σ_{HV}, water shows a wavy shape as incidence angle increases. In the first subswath, water shows larger variation compared to the rest, but generally σ_{HV} is not as obviously affected by incidence angle. With this special pattern in HV in Sentinel-1 imagery, previous methods tested on ScanSAR images, for instance, IRGS [23] which performs initial
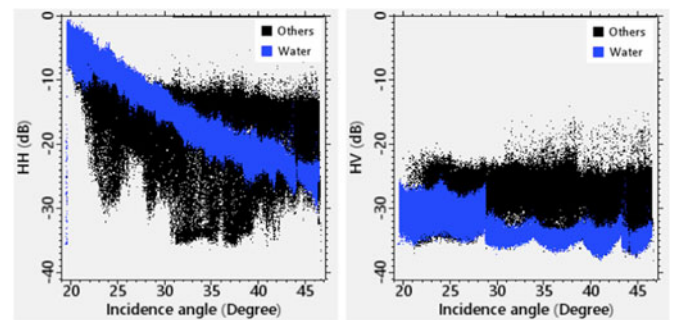

Fig. 6. Backscatter variation of water at different incidence angles in scene captured on January 18, 2016.

segmentation on HV bands, may not be applicable. Besides water, no obvious backscatter intensity differences were found in different incidence angles in most ice types. In addition, from the σ_{HV} values shown in Fig. 6, majority of the values were below noise floor in HV. Though the values would be unreliable, some patterns could still be observed.

However, despite some patterns could be found in HV, the noise in the first subswath brings too many uncertainties as the backscatter values greatly constrains the identification of different ice types and water. Attempts have been made on applying the proposed workflow on full Sentinel-1 scenes, but the segmentation results were not desirable, as patterns of the noise rather than ice were selected. As a result, removing the first subswath would be a compromise before ESA produces an official noise reduction solution.

B. Training Samples

The training samples were selected from four scenes: January 4, January 18, March 16, and April 2, 2016, with reference to

TABLE IV
MEAN σ_0 OF TRAINING SAMPLES

Dates	Jan 4, 2016		Jan 18, 2016		Mar 16, 2016		Apr 2, 2016	
	σ_{HH}	σ_{HV}	σ_{HH}	σ_{HV}	σ_{HH}	σ_{HV}	σ_{HH}	σ_{HV}
New ice	−21.1	−33.8	−32.4	−33.9	−23.8	−34.4	NA	NA
Grey ice	−14.2	−27.7	−12.8	−23.5	−16.5	−29.4	−14.7	−29.1
Grey white ice	−9.9	−22.3	−15.6	−28.3	−11.1	−26.7	−13.6	−25.7
Thin first year ice	−10.3	−22.1	−19.9	−33.0	−21.8	−33.8	−10.1	−26.4
Medium first year ice	NA	NA	NA	NA	−10.2	−22.4	−11.6	−23.0
Water	−7.28	−25.7	−18.4	−34.4	−14.0	−33.8	−16.9	−29.1

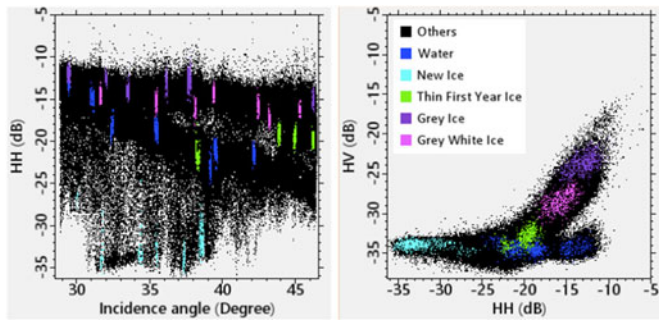The unit of σ_0 is dB; NA: not applicable

Fig. 7.   Scatterplots of training samples in the scene on January 18, 2016.

sea ice charts provided by the CIS. A total of 11 353 pixels were selected, and Table IV shows a summary of the training samples and mean values of σ_0. It can be observed that mean σ_{HH} and σ_{HV} values were very different in ice across different scenes and few patterns could be found. The ice types in different scenes are visually different as well only from σ_{HH} and σ_{HV}. Taking gray ice and gray white ice as an example, gray white ice is thicker than gray ice, which may result in higher response in both σ_{HH} and σ_{HV}, but in the scene on January 18, 2016, gray ice has both higher than σ_{HH} and σ_{HV} values. In addition, in water class, mean σ_{HH} values are highest among all the classes, except for the scene on April 2, 2016, where water locates at larger incidence. This may be a result of the incidence angle effect on water in σ_{HH}.

Two scatterplots are shown in Fig. 7 to show the separability of the training samples in this scene. From the scatterplot on the left, it can be observed that water has a large variation in σ_{HH} and values decrease as incidence angle increase, but other classes are relatively stable. No significant incidence angle dependence in σ_{HV} was found in all classes so that the scatterplot is not shown. From the scatterplot on the right, it can be concluded that water has the largest variation in both σ_{HH} and σ_{HV}, so that it could be mixed with ice classes just from a numerical perspective. New ice also showed some variation in σ_{HH} as some samples were selected in the first subswath. Gray ice and gray white ice have very similar σ_{HH} values, but some separability could be found in σ_{HV}. New ice, thin first year ice, and water are close in σ_{HH} values, while they could be easily separated in σ_{HV}, but

there are still some confusion majorly from the large variation of backscatter from water.

Considering the noise floor of Sentinel-1 EW mode is –22 dB, most of the σ_{HV} values in the images involved in this study were below −22 dB, making the backscatter from HV unreliable. However, the differences in σ_{HV} of different ice types can still be found both visually and from the selected samples shown in Table IV and Fig. 7. Thus, HV is expected to have contributions to image segmentation, and feature selection would justify whether HV is effective in identifying ice and water.

All training samples were selected manually based on our interpretation of the Sentinel-1 imagery with reference to sea ice charts. However, there are some uncertainties that may lead to errors or deficiency in the training samples. First, misinterpretation of ice charts may be one of the uncertainties. In ice charts, each polygon marked by egg codes usually contains two to three classes or even more, which increases the difficulty determining one certain ice type. Second, the strategy of sample selection may result in uncertainties. In order to efficiently select training samples, the strategy was to select "pure" pixels to capture the most significant features. As a result, some areas with ambiguity such as MIZ, where several ice types with different flow sizes occur, or where egg codes are difficult to understand, were avoided, and some patterns in these regions may be lost due to complexity and uncertainty. Finally, the scheme of labels may have an impact on the next steps. During training sample selection, labels were determined by referring to egg codes, and the six classes were: water, new ice, gray ice, gray white ice, thin first year ice, and medium first year ice. Among these classes, water alone may have a large variance in σ_0 values. Thus, in an unsupervised classification or segmentation framework, the ideal class scheme may not statistically fit in different images, and sometimes two or more subclasses would be more precise instead of one. However, the training samples were only used in feature selection, and the final segmentation result did not follow the labels in training samples, but a more reasonable class scheme might provide more effective feature selection results.

C. Feature Selection

During the three steps in the feature selection algorithm, three corresponding coefficients could be altered to achieve a smaller subset or to improve computational efficiency. Generally, by
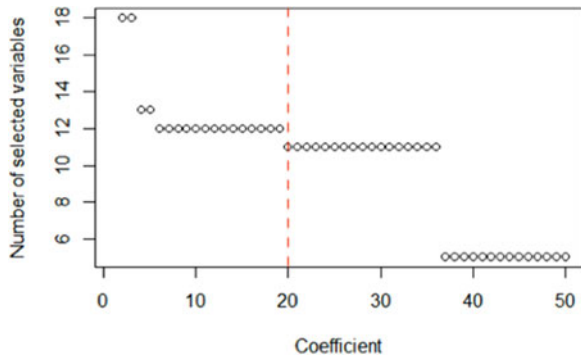
Fig. 8.   Selection of coefficient C_2 in interpretation step.

σ_{HH} Variance 11×11 step 1 σ_{HV} Contrast 25×25 step 5 σ_{HV} Mean 11×11 step 1

σ_{HV} Correlation 25×25 step 1 σ_{HH} Variance 25×25 step 5 σ_{HV} Dissimilarity 25×25 step 5

Fig. 9.   Six selected features of the scene on January 18, 2016.

choosing larger coefficients promote fewer features and faster selection speed. By conducting tests on how each of the three coefficients affect the number of features selected in each step, suitable coefficients could be selected. The first coefficient C_1 was used for setting the threshold for feature elimination, and it was multiplied by a minimum variable importance generated by CART function. Fewer features selected in the first step, the faster the interpretation step would be. A test using C_1 from 1 to 50 was conducted, and the number of remaining variables at the first step decrease as the coefficient increase. However, the selection from 1 to 50 did not change the result of the second step. The second coefficient C_2 in the interpretation step determines the smallest number of features that could produce comparable results comparing to using all features. A test was also conducted on how C_2 affects the number of feature selected during interpretation step. As shown in Fig. 8, two jumps could be found at 11 and 12 variables, and the two jumps indicate that the variables rank at 12th and 11th could be able to reduce E_{oob} significantly. Although another jump occurred when a relatively large coefficient (around 40) was adopted, it would result in only 5 variables selected in this step, which might be too aggressive using this setting. By referring to mean E_{oob} from the first 12 to 11 features, the error rate increase was minimal. Thus, 20 was selected as the coefficient in the interpretation step, and 11 variables were selected in this step. In terms of the third coefficient C_3 in the prediction step, it determines the final number of most effective features. During the test of C_3, 6 features were selected using C_3 from 1 to 30, and a coefficient larger than 30 led to 5 selected features or less. Since a coefficient of 30 would be considerably large, and using six features would result in significantly lower mean E_{oob} compared to 5, C_3 was kept as 1.

Fig. 9 shows the images of the six selected features in the scene on January 18, 2016. Among the six features, four of them were derived from σ_{HV} and the rest two were from σ_{HH}. The result indicated that HV could be more effective in separating different ice types and water regardless of its values are below the noise floor. From the appearance of the features in Fig. 9, features derived from σ_{HH} are still affected by incidence angle on water, while those derived from σ_{HV} have less influence of the banding effect and horizontal lines except "σ_{HV} Mean 11 × 11 step 1." Generally, the selected GLCM features from
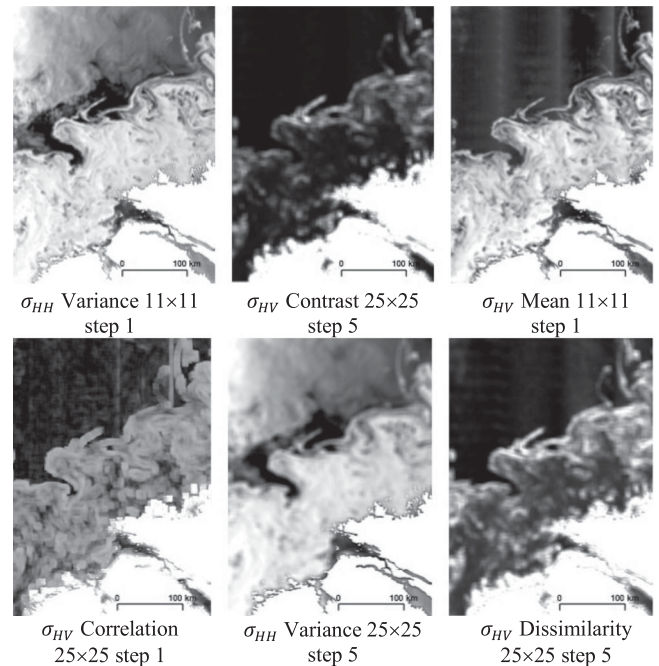
σ_{HV} could potentially mitigate the influence of noise in HV in Sentinel-1 imagery.

There are some uncertainties during the process of feature selection. First, the quality of training samples directly affects the results of feature selection. Since the selected training samples mostly from "pure" pixels and complicated regions were avoided, uniform characteristics could be expected within each class as well as surrounding pixels. However, some patterns may not be captured due to the strategy, so that these features may not be effective in those complicated regions. Second, the selected six features from the four scenes may not be the most significant features in each of them. Therefore, the selected six features may not lead to best performance in distinguishing different ice types and water in a certain scene, but an overall effective differentiation could be achieved. Third, the feature selection algorithm itself may not generate the optimal set of variables. In this forward searching algorithm, the searching order is based on the rank of importance produced by random forests. However, most significant individual variables may not eventually transform into an optimal variable set. Though the final six variables were able to produce a low error rate, they might not be the most effective set of 6, or a smaller set may exist with comparable performance.

D. Image Segmentation

Table III shows that most scenes were well segmented, and some scenes that are worth noting are discussed as follows.

The segmentation result of scene on March 16, 2016 is shown in Fig. 10, and the result achieved the user's accuracy of 14.2%, which is the lowest among all the results of 13 scenes. In addition, with the other two candidate label cost settings: 20 and 25, the user's accuracy did not increase, but the producer's
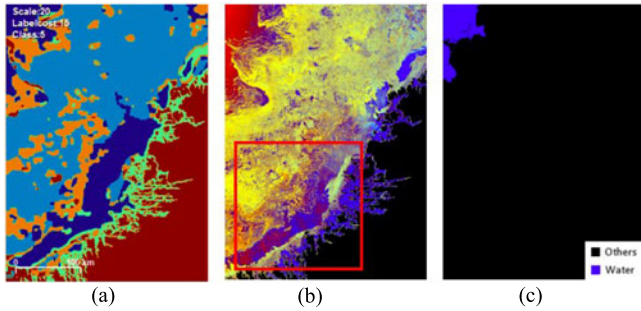
Fig. 10. Segmentation result of scene on March 16, 2016. (a) Segmentation result. (b) RGB (R: σ_{HH}, G: σ_{HV}, B: σ_{HH}/σ_{HV}) image of data. (c) Reference map of water.
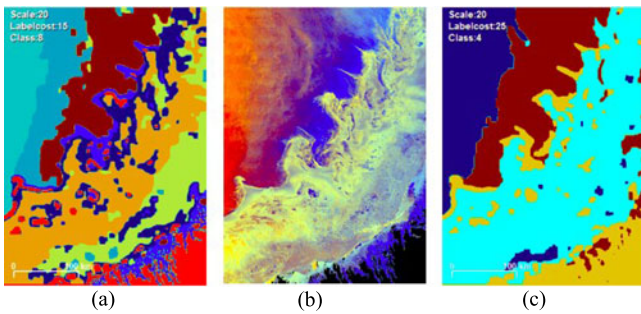

Fig. 11. Segmentation result of scene on March 23, 2016. (a) Segmentation result. (b) RGB (R: σ_{HH}, G: σ_{HV}, B: σ_{HH}/σ_{HV}) image of data. (c) Result with label cost 25.

Fig. 12. Segmentation result of scene on January 18, 2016. (a) Segmentation result. (b) RGB (R: σ_{HH}, G: σ_{HV}, B: $\frac{\sigma_{HH}}{\sigma_{HV}}$) image of data. (c) Reference map of water.

accuracy reached 88.2%, which is below average but still could be considered to be good. By comparing Figs. 10(a) and (b), which are the segmentation result and the RGB image, it could be found that the ice type in the red box was considered to be the same type with water. According to the ice chart on that date, the marked ice type was majorly new ice, which is the type that is difficult to separate from water. Though the water identification was not successful, the segmentation of water and new ice in this scene were relatively complete and accurate by visual comparison with the ice chart. Further tests on using different label cost settings did not provide satisfying results so that the mislabeling could be owed to the backscatter similarity between new ice and water. Therefore, the result shows that the discrimination between new ice and water needs to be improved to achieve an accurate segmentation result.

Among the segmentation results achieved using label cost of 15, the results of scene on March 23, 2016 resulted in the most number of labels. As Fig. 11(a) shows, a total of eight labels were produced by this segmentation algorithm, and three of them were presenting water. Both numbers are the highest among all the segmentation results listed in Table III. Though the large number of labels makes image interpretation unintuitive, user's accuracy, producer's accuracy, and overall accuracy reached 95.3%, 95.9%, and 97.0%, respectively, which could be considered a successful segmentation result. However, a better segmentation result could be achieved as the number of labels has the potential to be further reduced. By referring to the results shown in Table III, the number of class reduces as label cost increases to 20 and 25. As shown in Fig. 11(c), with the label
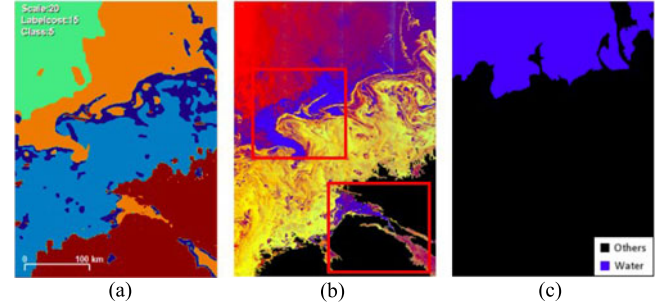
cost setting at 25, four total number of labels were produced by the segmentation algorithm, among which two classes represent water, but by further increasing label cost, unreasonable results were produced because of excessive merging. The label cost setting at 25 resulted in user's accuracy, producer's accuracy, and overall accuracy of 85.9%, 98.1%, and 93.8%, respectively. The larger label cost resulted in 10% decrease in user's accuracy since some ice areas were incorrectly merged with water at the right side of the image as shown in Fig. 10(c), but the reduction in overall accuracy was not that obvious. Therefore, though lower user's accuracy was generated with the large label cost setting, 25 may be considered a better choice in the scene on March 23, 2016 since it reduces number of labels significantly, enabling more intuitive interpretation of water contents in the image in the scope of the study.

Fig. 12 illustrates the segmentation result of the scene on January 18, 2016 where some ambiguities occur. The user's accuracy, producer's accuracy, and overall accuracy of this segmentation result were 61.8%, 99.8%, and 85.4%, respectively. The relatively low accuracy was mainly caused by the mislabeling in the two areas in red boxes as shown in Fig. 11(b). In the upper red box shows an area that might be new ice, as evidence could be seen in HH but not in HV (as shown previously in Fig. 4.). This region was labeled as ice in manual labeling and training sample selection, but it is questionable as the backscatter values were below the noise floor in both σ_{HH} and σ_{HV}. Another area shown in the lower red box shows the area close to the land, and the area would be ambiguous as low ice concentration was found in the ice chart. The regions close to land may have more uncertainties as the radar backscatter may be influenced by water depth below ice or the regions could be ice frozen to the ground. The selected six features were not able to successfully capture effective patterns to distinguish water and ice, and the algorithm could be further improved in these ambiguous regions.

The results of the selected parameter setting could generate results of close to 95% in the overall accuracy, but there are some identified uncertainties and limitations in the image segmentation process. First of all, according to the results listed in Table III and the discussions on specific images, it could be concluded that no parameter setting was perfect for all of the 15 images, or even most of them, not mentioning if the method

is applied in operational image segmentation. In evaluation of successfulness of water identification, the appropriate number of labels might vary. Even though the label cost settings could decrease the number of labels in each image, different data distribution, different ice types, and different incidence angles may result in different number of labels during the optimization process. Second, in this study, the selected features and the data model may not be optimal. Banding effect of Sentinel-1 imagery was mitigated by the selected features to some extent, but some parts of water may be similar to some ice classes in some cases. In addition, the features were selected using random forests, which are based on decision trees with arbitrary thresholds, but the data model used in segmentation optimization was different. Consequently, high classification results in features selection may not necessarily lead to satisfying labeling results. In terms of selection of data model, Euclidian distance was chosen, but it may not be the best choice with the selected features or the Sentinel-1 SAR imagery. Third, shortage of ground truth may result in biased training sample selection and accuracy assessment. Training sample selection based on the author's preferences may not have a complete representation of each class. In addition, the misinterpretation of ice chart could bring error and uncertainties in the selected samples. This kind of problem also exist when manually providing ground truth for water, especially at MIZ. The boundaries of water and ice are very smooth in ice charts, while the true boundaries are not. Since very thin ice may not be captured by C-band SAR imagery, these "ice-water" boundaries in this study could actually be the threshold of whether ice could be identified by the sensor. Considering the development or melting of ice is a gradual process, the threshold might not be easy to find. Last but not least, the segmentation algorithm has limited comparison with other popular image segmentation methods. The segmentation algorithm has the unique advantage of identifying number of labels with prior settings and the application in remote sensing is limited; thus, makes it difficult to compare with other methods. Even if the performance is not as good as other segmentation methods with predefined number of labels, it is still worth testing and developing.

VI. Conclusion

SAR imaging satellites has been approved as an important data source for sea ice monitoring in Canada. However, the limited revisiting time of a single SAR satellite such as RADARSAT-2 and large volume of manual work are two main challenges in providing high-quality sea ice map products timely and effectively. To deal with the two challenges, new data sources and automated image processing algorithms are in demand. This study proposed a semiautomated SAR image segmentation workflow, which has been tested on 15 scenes of Sentinel-1 SAR images in the study area of Labrador coast.

Three main objectives were designed in this study: 1) to determine the most important features in identifying different types of sea ice, 2) to examine the effectiveness of the proposed workflow, and 3) to evaluate the capability of Sentinel-1 SAR imagery in sea ice mapping. A prototype automated image segmentation method has been developed, and experimental results have derived some insights on the three objectives. The results are concluded as follows.

First, six GLCM features were selected as the most significant features in distinguishing water and different ice types. These six features are: σ_{HH} Variance 11×11 step 1, σ_{HV} Contrast 25×25 step 5, σ_{HV} Mean 11×11 step 1, σ_{HV} Correlation 25×25 step 1, σ_{HH} Variance 25×25 step 5, and σ_{HV} Dissimilarity 25×25 step 5. Second, the proposed method was able to segment the 13 test images into 3–8 classes with the selected parameter set, which may potentially provide a solution to determine the suitable number of labels. However, it was also found that one setting of parameters was not able to provide optimal results in all images. Third, in Sentinel-1 SAR imagery, incidence angle effects in copolarized bands of water and banding effects in cross-polarized bands may affect the interpretation of sea ice, but these effects were mitigated by the selected features and the segmentation algorithm. The overall accuracy of the tests reached 95% in distinguishing water in the images, and most errors comes from the similarity between water and thin ice types. After all, higher accuracies could be achieved when tuning parameters in individual images, but it may not be applicable if an automated method is needed.

In conclusion, Sentinel-1 SAR imagery is able to monitor sea ice conditions effectively, and the proposed workflow has the potential of developing into an automated image segmentation solution with further development. However, the selected features and the algorithm settings limit to the study area and the involving ice types, performance in other areas and other ice types needs further exploration to enable operational use of the proposed workflow.

In the future, more work and experiments will be done in the following areas: 1) ground truth with higher quality should be acquired to enable training and validation of more precise ice types, and automatic training and classification could be developed; 2) more features and data models should be tested to improve the proposed image segmentation workflow, and other areas with different ice regimes should be tested; 3) correction of thermal noise in Sentinel-1 data should be performed to evaluate the workflow on full scene images; 4) more products of Sentinel-1 could be utilized, and extra ancillary data should be integrated to sea ice mapping.

References

[1] C. Kinnard, C. M. Zdanowicz, D. A. Fisher, E. Isaksson, A. de Vernal, and L. G. Thompson, "Reconstructed changes in Arctic sea ice over the past 1,450 years," *Nature*, vol. 479, no. 7374, pp. 509–512, 2011.

[2] J. E. Walsh, F. Fetterer, S. J. Scott, and W. L. Chapman, "A database for depicting Arctic sea ice variations back to 1850," *Geograph. Rev.*, vol. 107, pp. 89–107, Jul. 2016.

[3] National Snow and Ice Data Centre, "Daily image update." 2016. [Online]. Available: http://nsidc.org/arcticseaicenews/2012/05/daily-image/, Accessed on: May 25, 2016.

[4] M. C. Serreze, M. M. Holland, and J. Stroeve, "Perspectives on the Arctic's shrinking sea-ice cover," *Science*, vol. 315, no. 5818, pp. 1533–1536, Mar. 2007.

[5] R. Kwok and D. Rothrock, "Decline in Arctic sea ice thickness from submarine and ICESat records: 1958–2008," *Geophys. Res. Lett.*, vol. 36, no. 15, pp. L15501-1–L15501-5, Aug. 2009.

[6] M. Wang and J. E. Overland, "A sea ice free summer Arctic within 30 years: An update from CMIP5 models," *Geophys. Res. Lett.*, vol. 39, no. 18, pp. 1–5, Sep. 2012.

[7] D. Notz and J. Stroeve, "Observed Arctic sea-ice loss directly follows anthropogenic CO_2 emission," *Science*, vol. 354, no. 6313, pp. 747–750, Nov. 2016.

[8] J. A. Screen and I. Simmonds, "The central role of diminishing sea ice in recent Arctic temperature amplification," *Nature*, vol. 464, no. 7293, pp. 1334–1337, 2010.

[9] J. Morison, *et al.*, "Changing Arctic ocean freshwater pathways," *Nature*, vol. 481, no. 7379, pp. 66–70, 2012.

[10] E. Post *et al.*, "Ecological consequences of sea-ice decline," *Science*, vol. 341, no. 6145, pp. 519–524, 2013.

[11] L. Xu and J. Li, "Mapping sea ice from satellite SAR Imagery," in *Monitoring and Modeling of Global Change: A Geomatics Perspective*. J. Li and X. Yang Eds. New York, NY, USA: Springer, 2015, pp. 113–136.

[12] Office of the Auditor General of Canada, "Marine navigation in the Canadian Arctic," Fall Rep. Commissioner Environ. Sustainable Dev., 2014.

[13] J. Karvonen, "Baltic Sea ice concentration estimation based on C-band dual-polarized SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5558–5566, Sep. 2014.

[14] R. Bintanja and F. M. Selten, "Future increases in Arctic precipitation linked to local evaporation and sea-ice retreat," *Nature*, vol. 509, no. 7501, pp. 479–82, 2014.

[15] F. De Zan and A. M. Guarnieri, "TOPSAR: Terrain observation by progressive scans," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2352–2360, Sep. 2006.

[16] D. Geudtner, R. Torres, P. Snoeij, M. Davidson, and B. Rommen, "Sentinel-1 system capabilities and applications," in *Proc. IEEE Geosci. Remote Sens. Symp.*, 2014, pp. 1457–1460.

[17] M. Arkett *et al.*, "Preparation by the Canadian Ice Service for the operational use of the RADARSAT Constellation Mission in their ice and oil spill monitoring programs," *Can. J. Remote Sens.*, vol. 41, no. 5, pp. 380–389, 2015.

[18] N. Y. Zakhvatkina, V. Y. Alexandrov, O. M. Johannessen, S. Sandven, and I. Y. Frolov, "Classification of sea ice types in ENVISAT synthetic aperture radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2587–2600, May 2013.

[19] R. Ressel, A. Frost, and S. Lehner, "A neural network-based classification for sea ice types on X-band SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 7, pp. 3672–3680, Jul. 2015.

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

[21] Q. Yu and D. A. Clausi, "IRGS: Image segmentation using edge penalties and region growing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2126–2139, Dec. 2008.

[22] S. Ochilov and D. A. Clausi, "Operational SAR sea-ice image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4397–4408, Nov. 2012.

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

[24] S. Leigh, Z. Wang, and D. A. Clausi, "Automated ice-water classification using dual polarization SAR satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5529–5539, Sep. 2014.

[25] F. Li, D. A. Clausi, L. Wang, and L. Xu, "A semi-supervised approach for ice-water classification using dual-polarization SAR satellite imagery," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 28–35.

[26] T. Kwon, J. Li, and A. Wong, "ETVOS: An enhanced total variation optimization segmentation approach for SAR sea-ice image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 925–934, Feb. 2013.

[27] J. Karvonen, "Baltic Sea ice concentration estimation based on C-band HH-polarized SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 6, pp. 1874–1884, Dec. 2012.

[28] L. Xu, D. A. Clausi, F. Li, and A. Wong, "Weakly supervised classification of remotely sensed imagery using label constraint and edge penalty," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1424–1436, Mar. 2017.

[29] M. Dabboor and M. Shokr, "A new likelihood ratio for supervised classification of fully polarimetric SAR data: An application for sea ice type mapping," *ISPRS J. Photogramm. Remote Sens.*, vol. 84, pp. 1–11, 2013.

[30] 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 Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4524–4533, Aug. 2016.

[31] N. G. Kasapoglu, "Sea ice concentration retrieval using composite ScanSAR features in a SAR data assimilation process," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2085–2089, Dec. 2014.

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

[33] L. Xu, J. Li, Y. Shu, and J. Peng, "SAR image denoising via clustering-based principal component analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6858–6869, Nov. 2014.

[34] 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 Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3188–3199, Jun. 2016.

[35] L. Xu, J. Li, A. Wong, and C. Wang, "A KPCA texture feature model for efficient segmentation of RADARSAT-2 SAR sea ice imagery," *Int. J. Remote Sens.*, vol. 35, no. 13, pp. 5053–5072, 2014.

[36] R. Genuer, J. Poggi, and C. Tuleau-Malot, "Variable selection using random forests," *Pattern Recognit. Lett.*, vol. 31, no. 14, pp. 2225–2236, Oct. 2010.

[37] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[38] A. Delong, A. Osokin, H. N. Isack, and Y. Boykov, "Fast approximate energy minimization with label costs," *Int. J. Comput. Vis.*, vol. 96, no. 1, pp. 1–27, Jan. 2012.

[39] K. Xu, W. Yang, G. Liu, and H. Sun, "Unsupervised satellite image classification using Markov field topic model," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 1, pp. 130–134, Jan. 2013.

[40] P. Siva and A. Wong, "URC: Unsupervised regional clustering of remote sensing imagery," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2014, pp. 4938–4941.

[41] J. Yan, J. Shan, and W. Jiang, "A global optimization approach to roof segmentation from airborne lidar point clouds," *ISPRS J. Photogramm. Remote Sens.*, vol. 94, pp. 183–193, Aug. 2014.

[42] R. Szeliski *et al.*, "A comparative study of energy minimization methods for Markov random fields with smoothness-based priors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 6, pp. 1068–1080, Jun. 2008.

[43] D. J. Cavalieri and C. L. Parkinson, "Arctic sea ice variability and trends, 1979-2010," *The Cryosphere*, vol. 6, no. 4, pp. 881–889, Aug. 2012.

[44] I. Fenty and P. Heimbach, "Coupled sea ice–ocean-state estimation in the Labrador Sea and Baffin Bay," *J. Phys. Oceanograph.*, vol. 43, no. 5, pp. 884–904, May 2013.

[45] B. Ellis and L. Brigham, "Arctic marine shipping assessment 2009 report," 2009, [Online]. Available: https://www.pmel.noaa.gov/arctic-zone/detect/documents/AMSA_2009_Report_2nd_print.pdf

[46] Canadian Coast Guard, "Ice navigation in Canadian waters," Ottawa, ON, Canada, 2012.

[47] European Space Agency, "Sentinel-1 sar user guides," 2016. [Online]. Available: https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar/acquisition-modes/extra-wide-swath, Accessed on: Sep. 25, 2016.

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

[49] J.-W. Park, A. A. Korosov, M. Babiker, S. Sandven, and J.-S. Won, "Efficient thermal noise removal for Sentinel-1 TOPSAR cross-polarization channel," *IEEE Trans. Geosci. Remote Sens.*, pp. 1555–1565, doi: 10.1109/TGRS.2017.2765248.

[50] J. Karvonen, "Baltic sea ice concentration estimation using Sentinel-1 SAR and AMSR2 microwave radiometer data," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2871–2883, May 2017.

[51] N. Miranda *et al.*, "The Sentinel-1 constellation mission performance," in *Proc. Int. Radar Conf.—Surveillance Safer World*, pp. 5541–5544, 2017.

[52] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man. Cybern.*, vol. 3, no. 6, pp. 610–621, Nov. 1973.

[53] D. G. Barber and E. F. LeDrew, "SAR sea ice discrimination using texture statistics: A multivariate approach," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 4, pp. 385–395, 1991.

[54] A. Zvoleff, "GLCM: Calculate textures from grey-level co-occurrence matrices (GLCMs)," 2016, Online. Available: https://cran.r-project.org/web/packages/glcm/index.html

[55] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

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

[57] R. Genuer, J.-M. Poggi, and C. Tuleau-Malot, "VSURF: An R package for variable selection using random forests," *The R J.*, vol. 7, no. 2, pp. 19–33, Dec. 2015.

[58] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Boca Raton, FL, USA: CRC Press, 1984.

[59] W. Tan, R. Liao, Y. Du, J. Lu, and J. Li, "Improving urban impervious surface classification by combining Landsat and PolSAR images: A case study in Kitchener-Waterloo, Ontario, Canada," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, pp. 1917–1920, 2015.

[60] O. Veksler and A. Delong, "GCoptimization—software for energy minimization with graph cuts," 2010, Online. Available: http://vision.csd.uwo.ca/code/

[61] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.*, vol. 37, no. 1, pp. 35–46, Jul. 1991.
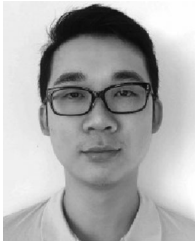
Weikai Tan received the B.Sc. degree from Nanjing University, Nanjing, China, in 2014 and the B.Sc. and M.Sc. degrees in geomatics from the University of Waterloo, Waterloo, ON, Canada, in 2014 and 2016, respectively.

His research interests include SAR image processing, sea ice mapping, PolSAR image analysis, land cover classification.

Jonathan Li (M'00–SM'11) received the Ph.D. degree in geomatics engineering from the University of Cape Town, Cape Town, South Africa, in 2000.

He is currently a Professor with both the Departments of Geography and Environmental Management as well as Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada. His main research interests comprise LiDAR and SAR data processing, machine learning, and remote sensing applications. He has coauthored more than 400 publications, including more than 170 refereed journal papers.

Dr. Li is currently an Associate Editor of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING and the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS.
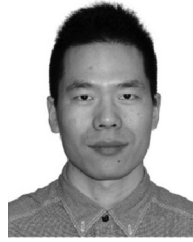
Linlin Xu (M'14) received the B.Eng. and M.Sc. degrees in geomatics engineering from the China University of Geosciences, Beijing, China, in 2007 and 2010, respectively, and the Ph.D. degree in geography from the University of Waterloo, Waterloo, ON, Canada, in 2014.

He was a Postdoctoral Research Fellow with the Department of Systems Design Engineering, University of Waterloo, Canada, from 2014 to 2017. He is currently an Associate Professor with the School of Land Science and Technology, China University of Geosciences, Beijing. His research interests include the areas of remote sensing image processing, especially hyperspectral and SAR image processing.

Michael A. Chapman received the Ph.D. degree in photogrammetry from Laval University, Québec City, QC, Canada.

He is currently a Professor with the Department of Civil Engineering, Ryerson University, Toronto, ON, Canada. Prior to joining Ryerson University, he was a Professor of Geomatics Engineering with the University of Calgary, Canada, for 18 years. He has authored or coauthored more than 160 technical articles. His research interests include algorithms and processing methodologies for airborne sensors using Global Navigation Satellite System (GNSS)/Inertial Measurement Unit, geometric processing of digital imagery in industrial environments, terrestrial imaging systems for transportation infrastructure mapping, and algorithms and processing strategies for biometrology applications.