Operational SAR Sea-Ice Image Classification

Shuhratchon Ochilov, Student Member, IEEE
David A. Clausi, Senior Member, IEEE

***Abstract*—Thousands of spaceborne synthetic aperture radar (SAR) sea-ice images are systematically processed every year in support of operational activities such as ship navigation and environmental monitoring. An automated approach that generates a pixel-level sea-ice image classification is required since manual pixel-level classification is not feasible. Currently, using a standardized approach, trained ice analysts manually segment full SAR scenes into smaller polygons to record ice types and concentrations. Using this data, pixel-level classification can be achieved by initial unsupervised segmentation of each polygon followed by automatic sea-ice labeling of the full scene. A fully automated Markov random field model that is used to assign labels to all segmented regions in the full scene has been designed and implemented. This approach is the first known successful end-to-end process for operational SAR sea-ice image classification. In addition, a novel performance evaluation framework has been developed to validate the segmentation and labeling of SAR sea-ice images. A trained sea-ice expert has conducted an arms length evaluation using this framework to generate a set of full scene reference images used for testing. Testing demonstrates operational success of the labeling approach.**

***Index Terms*—Synthetic aperture radar, sea-ice, image classification, unsupervised segmentation, Markov random field (MRF).**

I. Introduction

Recent developments in polar regions and the trend of global warming provide clear evidence that the interest in sea-ice monitoring applications will increase. Remote sensing, particularly satellite SAR imaging, is used to study behavior and change of sea-ice. Subsequently, fast processing and interpretation of large volumes of wide expanse satellite data is required. Operationally, SAR images are manually interpreted by assigning ice types and their concentrations to large regions. An automatic procedure that assigns ice type labels to pixels is preferable to generate pixel-level ice concentration maps.

The authors are with the Vision and Image Processing (VIP) Research Lab, Department of Systems Design Engineering, University of Waterloo, 200 University Ave. West, Waterloo, Ontario, Canada, N2L 3G1. Tel.: +1 519 888 4567 x32604. Fax: +1 519 746 4791. E-mail: dclausi@uwaterloo.ca.

The drawback of region-based interpretation is that the ice concentrations are often inaccurate and such data can not always be effectively used. For example, a region may have 30multi-year ice, but the user can not pinpoint the multi-year ice location. Such information would be useful for mathematical climate models, route planning and for ship navigation. Hence, automated interpretation of SAR sea-ice images would be invaluable for organizations performing sea-ice interpretation operationally or conducting research in this field. This paper describes an end-to-end operational process which provides pixel-level classification of full SAR sea-ice imagery.

The processing flow of SAR sea-ice imagery and the depiction of the problem is shown in Fig. 1. The first step is an image acquisition by a SAR satellite (Fig. 1a). Later, an ice analyst manually divides the image into large ”polygon” regions (Fig. 1b) and reports the ice types and their estimated concentrations for each polygon using an ”egg code”. The egg code, named after its oval shape, is the World Meteorological Organization (WMO) [1] standard which lists ice types and their concentrations without spatially identifying the location of each ice type in the polygon region. To automatically interpret images based on egg code data, unsupervised segmentation can be performed on each polygon independently (Fig. 1c). The challenge and focus of this paper is to automatically assign a sea-ice label to each segmented region in each polygon across the entire SAR scene (Fig. 1d).

Other research has considered the classical classification approach using training test samples [2]-[3]. Here, in a similar starting point as [4], we perform the classification using the polygon data. Each polygon is automatically segmented into disjoint regions and a global approach using all polygons is implemented to label the regions. This leads to an automatic and unsupervised algorithm for at-resolution classification of SAR sea-ice imagery.

The labeling technique uniquely models the spatial relationship of regions between the poly- gons in the form of a neighborhood system embedded in a Markov random field (MRF) [5] framework. The statistical and spatial relationship models are then combined in a Bayesian framework where the region labeling is for-

mulated as a global energy minimization problem. The mathematical model is described in Section III and its implementation in Section IV.

A consistent challenge in sea-ice classification research is the lack of validated full scene data. Field studies can not feasibly be performed for a single scene where the swath is 500km over hazardous ocean regions. The most effective validation is performed by trained ice experts. As such, we have created a user-interactive performance evaluation framework (Section V) whereby an expert third party can validate both the segmentation and labeling algorithmic outcomes. This is used to both evaluate the herein developed techniques and generate validated reference images. The reference database consists of three operational full-scene SAR sea-ice images which have been used to assess the performance of the proposed technique (Section VI).

II. BACKGROUND

A. SAR sea-ice remote sensing and processing

SAR is an active microwave sensor, which can capture images day and night irrespective of fog and clouds. These features makes SAR crucial for sensing polar regions. In the past two decades with the launch of RADARSAT-1,2, Japanese Earth Resource Satellite (JERS)-1,2, European Remote Sensing (ERS)-1,2, Environmental Satellite (ENVISAT) and others [6], there has been a shift from aerial platforms towards the use of spaceborne platforms for capturing SAR sea-ice imagery because of reduced costs.

RADARSAT-1,2 are commercial SAR satellites launched in 1995 and 2007 with the primary goal of monitoring and managing sea-ice. RADARSAT-1 is operational but its estimated lifespan has been exceeded. To ensure continuity of service, the more advanced satellite RADARSAT- 2 was launched. The primary user of RADARSAT SAR images has been the Canadian Ice Service (CIS) [7] processing approximately 4000 SAR sea-ice images annually [8]. CIS is an organization responsible for providing sea-ice conditions of Canadian shores and seas, in a timely manner, to support safe maritime operations [7]. Both satellites operate in 5.3GHz C-band. RADARSAT-1 images [9] acquired in ScanSAR mode are the main source of SAR sea-ice imagery provided by CIS and used in this research. In ScanSAR mode the swath reaches the 500 km [10] range at resolution of 50 m. CIS 2x2 block averages SAR images for archival purposes to the image size of roughly 5000 by 5000 pixels. Using the RADARSAT-1,2 SAR sea-ice images and ancillary information from ships, aircrafts and meteorological sensors, a CIS analyst manually produces ice maps with egg codes and polygons similar to the map shown in Fig. 2. The standard sea-ice symbol representation along with descriptions are given in Table I.

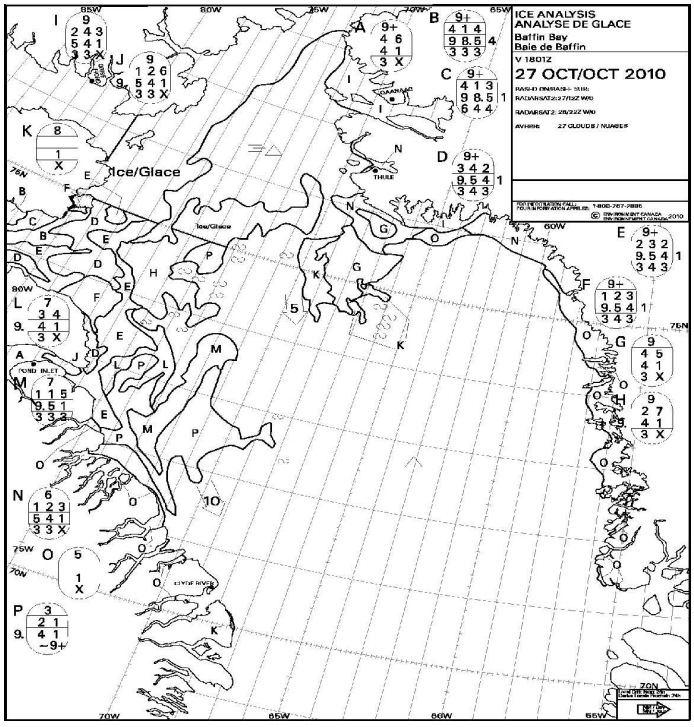


Fig. 2. The example of ice map provided by CIS (reduced to fit). Alphabet letters (A,B,C,...) associate egg code [1] with polygon regions.

B. Early Efforts

There is a history of supervised classification. For example, numerous studies [2],[3] require selection of training samples to extract class statistics and a priori knowledge of the class statistics is required for more recent methods [12],[13]. In this approach, the ice analyst cumbersomely discerns and selects pure ice samples from the image to train and test the classification. Such training data is subject to the bias of the individual analyst and does not consistently match the ice types across scenes and not even across a single scene. This variation is a result of the high inter-class and intra-class variability of the backscatter. The incidence angle, natural ice variability and environment are key contributors to nonstationarity in SAR sea-ice imagery. Hence, for operational use, techniques that classify based on thresholds [14], [15], [16] and statistics derived from training samples will not perform robustly. In reality, the human ice analyst differentiates ice types within the scene on a relative basis.

Published research [11], [17], [18] exists that segments SAR sea-ice scenes using an unsupervised approach.

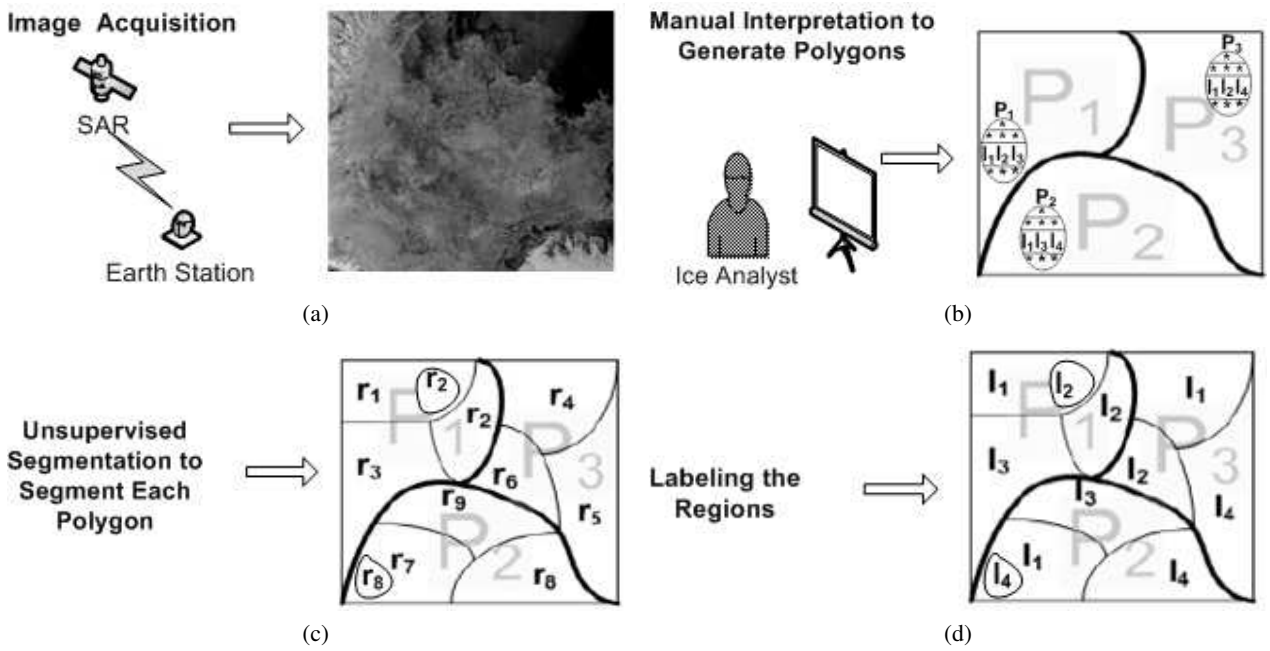


Fig. 1. Classification of full scene operational SAR sea-ice images. r, P, l refer to region, polygon and sea-ice label. (a) SAR image acquired by satellites. (b) Image manually divided into polygons with appropriate egg code data. (c) Image with every polygon automatically segmented into n_r regions using IRGS [11]. (d) Image with every region automatically labeled with sea-ice type. This paper focuses on (d), the labeling problem.

TABLE I
THE DESCRIPTION OF STAGE OF SEA-ICE DEVELOPMENT S_0, S_a, S_b, S_c, S_d [1]

Stage of development	Thickness(cm)	Ice-type code
New ice	< 10	1
Young ice	10 – 30	3
Grey ice	10 – 15	4
Grey-white ice	15 – 30	5
First-year ice	30	6
Thin first-year ice	30 – 70	7
Medium first-year ice	70 – 120	1.
Thick first-year ice	> 120	4.
Old ice		7.
Second-year ice		8.
Multi-year ice		9.
Fast-ice		▲●
Undetermined ice		X

Among different techniques, IRGS [11] (Iterative Region Growing with Semantics) has shown robust performance segmenting both operational SAR sea-ice and general purpose imagery. IRGS has been successfully tested and validated by CIS personnel. IRGS has been integrated into a software system called MAGIC [17] to use CIS source data and enable consistent testing in an easy-to-use GUI framework. For these reasons, IRGS has been used as the segmentation algorithm in this paper. Briefly, IRGS works by first breaking the scene into small uniform regions using a watershed segmentation algorithm. Then, each region is assigned to a class based on a Markov random field model. Adjacent regions that have the same assigned class are greedily merged until the system energy can not be decreased further. These two steps (class assignment and merging) are iterated until merging cannot be performed further.

Despite progress in the SAR sea-ice segmentation field, limited research has been performed in ice type labeling. Existing classification techniques [15], [19] make use of unsupervised segmentation as their initial step but label the segmented images based on training statistics which makes these approaches potentially not robust in an operational environment. An initial attempt to avoid the training step [20] used a logical assignment of ice labels by cognitive reasoning and class statistics, but this methods outcome is dependent on the ordering of the polygons. However, labeling is preferred to be optimal over the whole SAR scene, and can be performed automatically by the best fit modeling of information from all polygons.

III. SEA-ICE LABELING MODEL

A. Graph Model and Definitions

Let S denote a discrete 2D rectangular image space of $M \times N$ pixels. Each pixel s represents a σ° [21] value. Consider representing the image space S as n_r closed regions $r = \{r1, r2, ..., r_{n_r}\}$ having boundaries ∂. The polygon $P = \{P1, P2...P_{n_p}\}$ is a higher structure defined on image space S where each P_q consists of one or more regions $\{ri, \dots\} \subseteq r$ (see Fig. 1c). Assuming

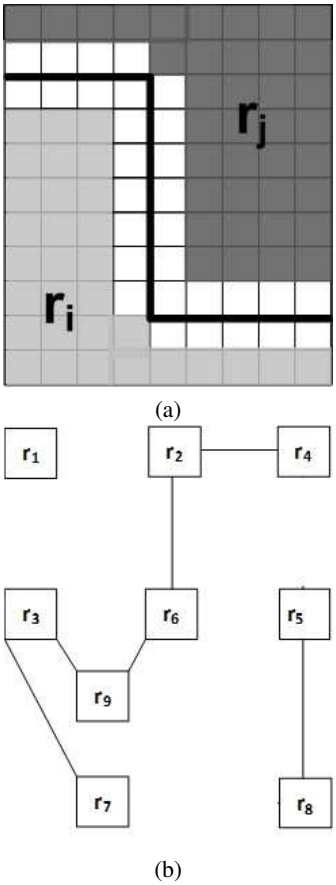


Fig. 3. (a)The boundary is the set of all white sites connecting two regions r_i, r_j based on first-order neighborhood system [22]. The bold black line is an edge separating two regions. (b)Graphical representation of Fig. 1c.

all polygons have been segmented, the segmented result $X = \{Xr \mid r \in S\}$ represents the random field defined on S, where X_r is the random variable representing all pixels $s \in ri$ in that region. Suppose there are n_l different sea-ice classes in X. Let $Z = \{Zr \mid r \in S\}$ be another random field defined on S where each discrete valued random variable Z_r , has value in $\{1, ...n_l\}$, represents the sea-ice label to which all pixels $s \in r_i$ in that region belong. Definitions essential to the proposed labeling model are given next.

Definition 1: Suppose the realization of Z, X, is $z = \{z_r \mid r \in S\}$ and $x = \{x_r \mid r \in S\}$, then the sea-ice labeling problem can be formulated as an estimation of z from x:

$$L : \Big\{ \{x_r \mid r \in S\} \quad \rightarrow \{z_r \mid r \in S\} \tag{1}$$

with the condition that the labeling realization $\{z_r \mid r \in P_q\}$ of some polygon Pq is constrained to one of the permutations of sea-ice labels assigned for that polygon.

Definition 2: The boundary ∂ is the set of all neighbor sites between regions. The neighborhood of sites is determined based on first-order neighborhood system [22] defined on image space S. If at least one such site exists, the regions are said to have a boundary between them as depicted in Fig. 3a.

Definition 3: Two regions r_i, r_j are neighbors if and only if the regions do not belong to same polygon $r_i, r_j \notin P_q$ for some polygon P_q and a boundary exists between them. In Fig. 1c r_2 and r_4 are neighbors but r_1 and r_2 are not. The graphical representation in Fig. 3b also shows this where the boxes each represent a region and the lines connecting the boxes define a neighbor relationship. N_{r_i} is the neighborhood of a region r_i comprising all regions r_j for which r_i, r_j are neighbors and has symmetrical relationship $r_j \in N_{r_i} \Leftrightarrow r_i \in N_{r_j}$. Neighborhood system N_r is a set of all neighborhoods.

Definition 4: Clique c is a single region or a subset of regions for which every pair of regions are neighbors and the set of all cliques is $C = [c \mid c \subset N_r]$ [22].

Definition 5: The random field S is an MRF with respect to the neighborhood system N_r if

$$p(z_r \mid S, r, s \notin S) = p(z_r \in N_r) \tag{2}$$

stating that the probability of some realization of the region given the whole graph is equal to probability given only neighbors. This property is also known as the Markovianity principle and its main advantage is the decomposition of large problems into smaller decorrelated and conditionally independent ones [23].

Theorem 1: The Hammersley-Clifford theorem states that if S is an MRF with respect to N_r and $p(z_r) > 0$ then S is a Gibbs random field (GRF) with respect to N_r where $p(z_r)$ can be expressed as [22] pp. 28.

$$p(z_r) = \frac{1}{Z}\exp\Big\{ -\frac{1}{T}\sum_{c\in C} V_c(x_r)\Big\} \tag{3}$$

where Z is the partition function that normalizes the distribution, T is the temperature, $V_c(x_r)$ is the clique potential and the negative of exponential $\sum_{c\in C} V_c(x_r)$ is called clique energy E. Theorem 1 is the framework for computing the probability of an MRF [24].

B. Sea-Ice Labeling

Labeling is performed by MAP (maximum a priori) estimation, derived from Bayesian theory, which maximizes the a posteriori probability:

$$p(Z = x \mid X = x) = \frac{p(X = x \mid Z = z)p(Z = z)}{\sum p(X = x \mid Z = z)p(Z = z)} \tag{4}$$

which states that a posteriori probability $p(Z = x \mid X = x)$ is equal to the product of likelihood $p(X = x \mid Z = z)$ and a priori probability $p(Z = z)$, divided by normalizing

constant [25]. The optimal estimator MAP can be given as:

$$\arg \max_{\{z_r | r \in S\}} p(Z = z \mid X = x) \propto$$
$$\arg \max_{\{z_r | r \in S\}} p(X = x \mid Z = z) p(Z = z) \tag{5}$$

where the a priori term $p(Z = z)$ is modeled as a Gibbs distribution (Eq. 3) and the feature model $p(X = x \mid Z = z)$ generally uses a Gaussian mixture model (GMM). If the Gibbs and Gaussian probability density functions are substituted into (Eq. 5), the maximization can be converted to a minimization by taking the negative of the exponentials. The minimization form is also known as the energy and the whole sea-ice labeling problem can be solved by energy minimization. Suppose the single clique and pairwise clique energy contributing from the feature and a priori models are E_f and E_p, respectively, then sea-ice labeling can be accomplished by solving the following optimization problem:

$$\arg \min_{\{z_r | r \in S\}} E_{total} \tag{6}$$

where,

$$E_{total} = \alpha E_f + \beta E_p \tag{7}$$

is the total energy. α and β are weighting parameters described in Section IV.

1) Feature model E_f: The statistical nature of SAR images indicates that the amplitude of the scattered signal is gamma distributed, however, in-house testing and published research [26] indicate that modeling classes in the feature space as Gaussian produces acceptable results. Thus, for simplicity, classes are assumed to be Gaussian and the energy of a single region is:

$$E(r_i) = \frac{k_{r_i}}{2} \ln 2\pi\sigma^2_{l_{r_i}} + \frac{k_{r_i}}{2\sigma^2_{l_{r_i}}} \left[(u_{r_i} - u_{l_{r_i}})^2 + \sigma^2_{r_i} \right] \tag{8}$$

where parameters u_{r_i} ,$\sigma^2_{r_i}$,k_{r_i} are the mean, covariance and the number of pixels of region r_i. In any realization of Y all r_i are assigned a sea-ice label, the $u_{l_{r_i}}$, $\sigma^2_{l_{r_i}}$ are the estimates of all r_i with the same label l_{r_i}. Equation (8) derived in Appendix A is computationally important since it calculates the energy of a region in a closed form using the region's statistics without iterating over the region's pixels. The total energy E_f over all regions is:

$$E_f = \sum_{i=1}^{n_r} E(r_i) \tag{9}$$

2) A priori term E_p: The regional representation of the standard multilevel logistic model (MLL) [22] can be defined as the energy of the Gibbs distribution with zero single node clique energy and pairwise clique energy given as:

$$E_p = \sum_{i=1}^{n_r} \left[\sum_{r_j \in N_{r_i}} \delta(l_{r_i}, l_{r_j}) \right] \tag{10}$$

where l_{r_i} and l_{r_j} are the labels assigned to r_i, r_j. The term $\delta(l_{r_i}, l_{r_j})$ is the Dirac delta function which equals 1 when $l_{r_i} = l_{r_j}$ and 0 otherwise. (Eq. 10) favors the configurations of neighboring regions having the same label. Such an a priori model is not suitable to model polygon interactions in full SAR sea-ice scenes. Polygons can be generated by an analyst to divide a large geographical region with the same ice type into smaller regions or to separate regions with different ice types, therefore two neighboring regions are not necessarily always of the same ice type. To accommodate such an interaction, here an edge penalty term has been incorporated into (Eq. 10).

$$E_p = \sum_{i=1}^{n_r} \left[\sum_{r_j \in N_{r_i}} g(\nabla_{r_i, r_j}) \delta(l_{r_i}, l_{r_j}) \right] \tag{11}$$

where

$$g(\nabla_{r_i, r_j} = 1 - \nabla_{r_i, r_j} \tag{12}$$

is the edge penalty term with edge strength ∇_{r_i, r_j} being the first order difference operator [27].Edge penalty increases if the edge strength is small, thereby, (Eq. 11) dictates that neighboring regions with lower edge strength are likely to be the same ice type.

IV. IMPLEMENTATION SCHEME

A. Optimization Scheme

A standard approach is used to find an optimal solution. Given (Eq. 9) and (Eq. 11) the combination of simulated annealing (SA) [28] and Metropolis sampling [29] has been applied using a common temperature schedule [30].

There are four parameters to be estimated: α, β, $u_{l_{r_i}}$ and $\sigma^2_{l_{r_i}} \forall i$ where r_i has been assigned the same label. As an unsupervised algorithm, the expectation-maximization (EM) [31], [32] algorithm can be used for estimating the class mean and covariance over a full SAR sea-ice scene. EM is suitable for maximum likelihood

estimation of feature parameters of incomplete data [31] and the convergence of the EM algorithm is known [33].

Here, β is set to one and the parameter α needs to be estimated accordingly. In conventional MLL models [22], α is a constant that can lead to solution divergence in early iterations due to too much weighting of the a priori model. To deal with this problem, the weight of the feature model can vary with each iteration [34] to improve performance. As such, the parameter α can be a function of iteration θ,

$$\alpha(\theta) = c_1 0.9^{\theta} + c_2 \tag{13}$$

where c_1, c_2 are constants equal to 0.1 [34]. Equation (13) monotonically decreases the parameter α with each iteration.

B. Algorithm Flow and Computational Efficiency

The algorithm is presented in seven modules as shown in Fig. 4 with each step defined below. While dealing with the typical large SAR sea-ice images the computational demands should be addressed for operational requirements.

1) The egg code data, the unsupervised segmentation result and the original calibrated SAR sea-ice image are loaded into memory. Therefore, the memory of the computing platform has to be sufficient for handling at least two full SAR sea-ice scenes totalling 800 MBytes.
2) The neighborhood is obtained using the spatial relationship of regions as specified in Fig. 3b and can be stored in a region adjacency graph (RAG) data structure. Also the edge penalty between every neighboring region $g(r_i, r_j)$ is computed. Both processes require the SAR sea-ice image in raster format with the computational complexity related to number of the image pixels $O(MN)$. Hence, in the proposed technique, edge penalty calculation in (Eq. 12) is computed once and stored as the regional edge penalty used for all iterations.
3) Only σ° is considered as a feature with associated mean ur_i, variance $\sigma^2_{r_i}$ and number of samples k_{r_i} stored for each region. These three values are sufficient to calculate the energy for r_i (Eq. 8). The computation of feature statistics is the last intensive processing step and has the complexity $O(MN)$.
4) The sea-ice labels are randomly assigned to regions as per Definition 1 to obtain the initial labeling realization Z.
5) Using the u_{l_r} and $\sigma^2_{l_{r_i}}$ from the previous step, (Eq. 7) is computed in E-step, which requires a pixel-based processing if not optimized. To accommodate this situation, a generic region-based formula is derived (Appendix A). This can turn hours of computational processing into a few seconds by reducing complexity from $O(\theta n_p MN)$ to $O(\theta n_r n_p + MN)$. During the minimization of (Eq. 7) with SA and Metropolis sampling the labeling refinement is accepted or rejected in the E-step using,

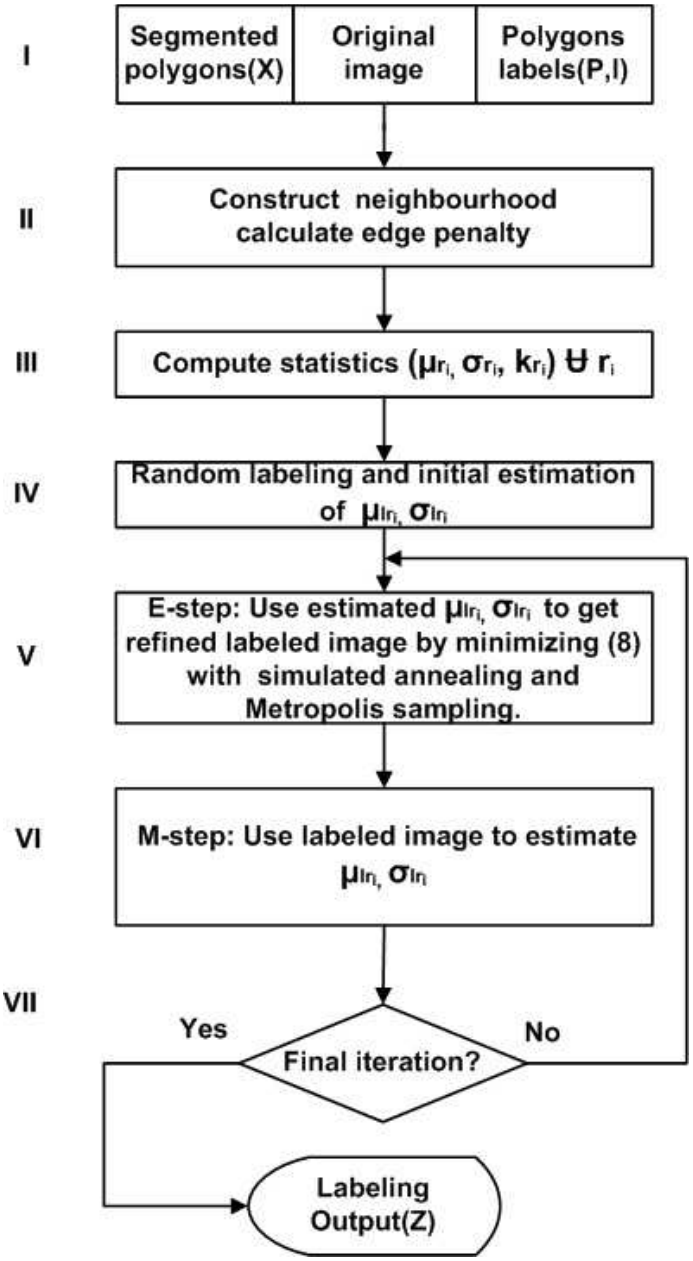


Fig. 4. Flowchart of labeling algorithm that performs Fig. 1. Indicated steps are discussed in text.

$$\rho > \exp\left[\frac{E_1 - E_2}{T}\right] \tag{14}$$

where ρ is the random number in the range

$[0, \ldots 1]$ drawn from a uniform distribution. E_1 and E_2 are the energies of current and potential labeled images. The stochastic nature of SA and Metropolis sampling avoids local minima by allowing both favorable and unfavorable changes to be accepted based on a temperature schedule [30].

6) Using the labeled image from the previous step, as part of the M-step, $u_{l_{r_i}}$ and $\sigma^2_{l_{r_i}}$ $\forall i$ are estimated by iteratively building the statistics from unions of regions $r_i j = ri \cup rj$ having the same sea-ice label. The formulation to achieve this task can be given as in [25]:

$$\begin{aligned} k_{r_{ij}} &= k_{r_i} + k_{r_j} \\ ur_{ij} &= \frac{u_{r_i} k_{r_i} + u_{r_j} k_j}{k_{ij}} \\ \sigma r_{ij} &= \frac{\sigma_{r_i} k_{r_i} + \sigma_{r_j} k_{r_j}}{k_{r_{ij}}} + \\ &\quad \frac{(u_{r_i} - u_{r_j})(u_{r_i} - u_{r_j})^t k_{r_i} k_{r_j}}{k^2_{r_{ij}}} \end{aligned} \tag{15}$$

where $k_{r_{ij}}$, ur_{ij} and σr_{ij} are the number of pixels, mean and variance after two regions are combined. The same process is followed for all regions r_i to obtain final parameters.

7) After energy oscillation at initial iterations the system is cooled down using a temperature schedule [30] to settle at global minima. The algorithm always converges well before reaching the set 100 iterations.

V. PERFORMANCE EVALUATION

A. Motivation

One of the main obstacles in developing algorithms for sea-ice interpretation is the lack of reference data. Fully validated field ground truth for the operational SAR sea-ice image is not available. For validation, one would have to perform field sampling of the sea-ice on site across 500km by 500km region during the SAR satellite overpass. Due to the logistical impossibility of such a validation exercise, we instead rely on the decades of CIS experience and know-how for interpreting SAR imagery. To systematically utilize CIS experience, a key contribution of this paper is a performance evaluation framework that has been implemented to guide an expert to generate a reference image. The evaluator (T. Zagon), a senior ice analyst with years of experience analyzing SAR sea ice images, has used this framework to generate reference images used for testing in this paper. To assess these images, the evaluator used as much time as necessary and accessed ancillary data to create a high level of confidence with the validation.

B. Images

The RADARSAT-1 HH images were selected deliberately by CIS to represent challenging examples for automatic classification and these were calibrated to σ^o. Table II provides summary information for each image and the geographical locations of images in the Arctic map as shown in Fig. 7. All the images are captured during freeze up which is especially challenging since the sea-ice properties are in transition.

C. Framework

Fig. 5 shows the flow of the algorithm framework. The framework is built within MAGIC [17]. Due to operational time constraints, it is challenging for an analyst to produce accurate data, and therefore the evaluation framework gives an opportunity to revisit the maps and eliminate the following common problems found with egg code data.

1) The number of classes provided per polygon might be incorrect. Ice types located near the polygon boundaries are not easily identifiable and might not be recognized in the egg code. The segmentation and labeling processes require the correct number of classes.
2) Ice analysts are generally biased towards assigning thicker ice types in polygons and overestimating thicker ice type concentrations. This is due to erring on the side of caution with regards to providing products for ship routing. This practice affects concentrations and ice typing for each polygon.

If required, the egg code parameters are adjusted by the evaluator. If the number of classes changes, the polygon is automatically resegmented otherwise the existing segmentation is used. The evaluator has to decide if the segmentation of the polygon is successful or if egg code parameters must be modified. If the segmentation fails, the evaluator discards the polygon from the reference image and enters an explanation for the failure. Eliminating any segmentation issues allows direct focus on evaluating the sea-ice labeling performance. The segmentation may fail in some exceptional scenarios where the discrimination depends on characteristics other than σ°. Apart from feature failures, the evaluator can discard a polygon if a portion of a land is included inside the polygon. This can happen in regions of no active ship navigation and such polygons need to be excluded to avoid erroneous classification.

For these CIS-provided data sets, the segmentation usually succeeds and then the evaluator assigns a class label to each region which acts as a ground truth. A

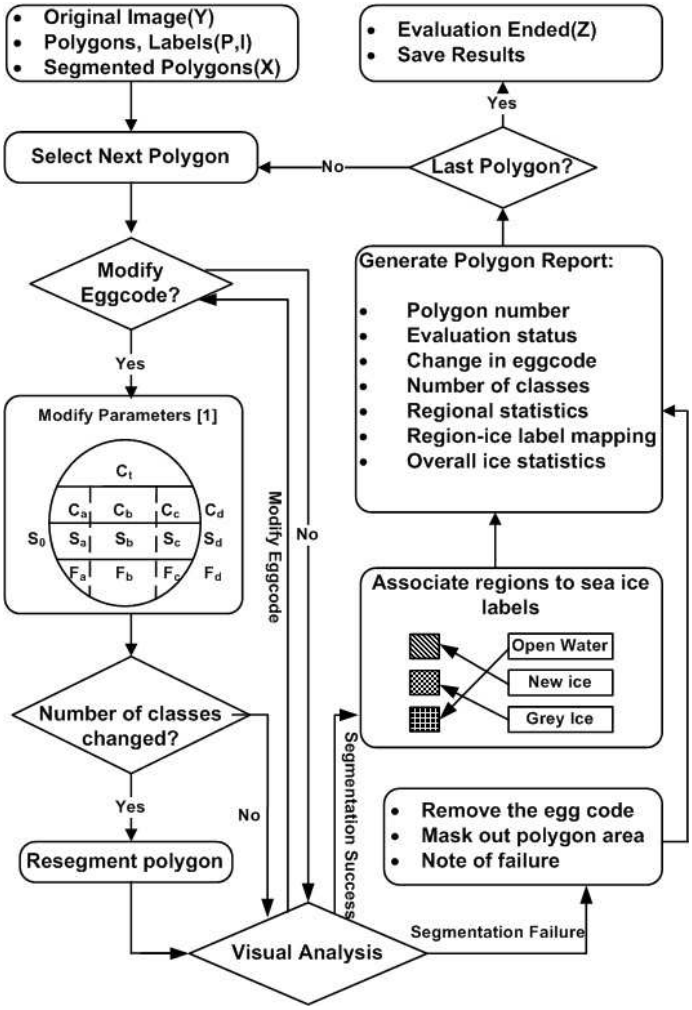


Fig. 5. This performance evaluation framework is integrated into MAGIC [17] and is designed for a person to generate full scene classified pixel-level data. The user can modify egg code parameters, resegment a polygon with new parameters and discard the polygon if segmentation fails. For successfully segmented polygons, the user manually assigns a sea-ice label to each region to create the reference database.

detailed report is generated which includes polygon information as well as all the steps the evaluator followed. Each polygon is evaluated in this manner until the full reference image is obtained.

D. Results of Performance Evaluation

Table II shows the results of the performance evaluation for each image including how many polygon eggcodes were modified and how many polygons were resegmented, discarded, and labelled. The evaluator had to make many changes to the CIS-provided source data to produce accurate reference data.

Nearly 80% (38 of 48) of the polygons were modified due to errors in the analyst provided data, including errors in the the number of ice classes and the ice types. Changes to the number of ice classes required about half of the polygons to be resegmented.

Some polygons were discarded. For Image 1, one polygon (Fig. 6a), according to the evaluator, requires a fine grey and new ice segmentation. The other Image 1 polygon (Fig. 6b) is a complicated scene that the evaluator was not able to visually interpret. For Image 2, the three erroneous polygons each have overly smoothed segmentations and an example is shown in Fig. 6c. Note that all polygons are segmented using the same algorithmic parameters within IRGS. For Image 3, five polygons were discarded due to errors in the provided land maps where land boundaries were not properly provided and islands were excluded (Fig. 6d). The evaluator excluded the polygons to remove any ambiguity.

Table III summarizes the remaining successfully segmented polygons with corresponding ice types used in the next section for comparisons.

VI. CLASSIFICATION RESULTS AND DISCUSSIONS

A. Evaluation

The images evaluated in Section V have been used as a basis for performance analysis of the labeling technique. As a performance metric, accuracy is calculated as a percentage of correctly classified pixels and regions, and kappa coefficient (κ) is used as a means of classification agreement [35]:

$$\kappa = \frac{P(A) - P(E)}{1 - P(E)} \tag{16}$$

where $P(A)$ is the probability the model values are equal to the actual value and $P(E)$ is the expected probability of random agreement. A value of one for kappa coefficient means a statistically perfect classification while a 0 means that all values are randomly classified. It is reasonable to assume that $\kappa \geq 0.7$ indicates accurate statistical modeling. The accuracy and the kappa metric are calculated based on both pixel-level and region-level accuracy. Pixel accuracy is the percentage of pixels properly classified. As a result, the pixel accuracy involves the size of the regions, whereas, the region accuracy is irrespective of the region size.

B. Labeling of operational SAR sea-ice images

As stated in Section II-B, IRGS is the algorithm used for segmentation. An example of an IRGS segmentation is shown in Fig. 8b. This polygon is downsized 25 times to fit so details are lost in the rescaling. This shows a typical challenging segmentation of grey, grey-white and new ice types which is not feasible to segment manually

TABLE II
PERFORMANCE EVALUATION OF SAR SEA-ICE IMAGES.

	Image 1	Image 2	Image 3
Image Area	Baffin Bay	Baffin Bay	Gulf of Boothila
Date Required	Oct. 30/05	Oct. 18/05	Oct. 06/04
Modified eddcode	12	13	13
Resegmented polygons	6	7	7
Discarded polygons	2	3	5
Labeled polygons	12	13	13
Total polygons	14	16	18

TABLE III
FINAL METADATA FOR REFERENCE DATABASE. USING TABLE I, LABELS REFER TO OPEN WATER(W), NEW ICE (1), GREY ICE (4), GREY WHITE ICE (5), OLD ICE (7.), SECOND YEAR ICE(8.) AND MULTI-YEAR ICE (9.) AS PER WMO [1] STANDARD.

Polygon	Image 1	Image 2	Image 3
P_1	$\{5,4,1\}$	$\{9,5,1\}$	$\{7,W\}$
P_2	$\{4,1\}$	$\{4,1\}$	$\{7,W\}$
P_3	$\{5,4,1\}$	$\{9,5,4,1\}$	$\{8,1\}$
P_4	$\{9,5,4\}$	$\{9,1\}$	$\{8,1\}$
P_5	$\{4,1\}$	$\{4,1\}$	$\{9,5,4,1\}$
P_6	$\{4,1\}$	$\{5,4\}$	$\{9,5,4,W\}$
P_7	$\{5,4,1\}$	$\{5,4,1\}$	$\{9,8,1,W\}$
P_8	$\{4,1\}$	$\{1,W\}$	$\{9,4,1,W\}$
P_9	$\{4,1\}$	$\{4,1\}$	$\{4\}$
P_{10}	$\{9,5,4\}$	$\{9,5,4,1,W\}$	$\{4\}$
P_{11}	$\{4,1\}$	$\{4,1\}$	$\{4\}$
P_{12}	$\{4,1\}$	$\{9,5,4,1\}$	$\{5,4\}$
P_{13}	N/A	$\{5,4,1\}$	$\{8,4,W\}$

TABLE IV
PERFORMANCE OF PROPOSED SEA-ICE LABELING TECHNIQUE OF ACCURACY/KAPPA AND RATIO OF CORRECTLY LABELED/TOTAL REGIONS.

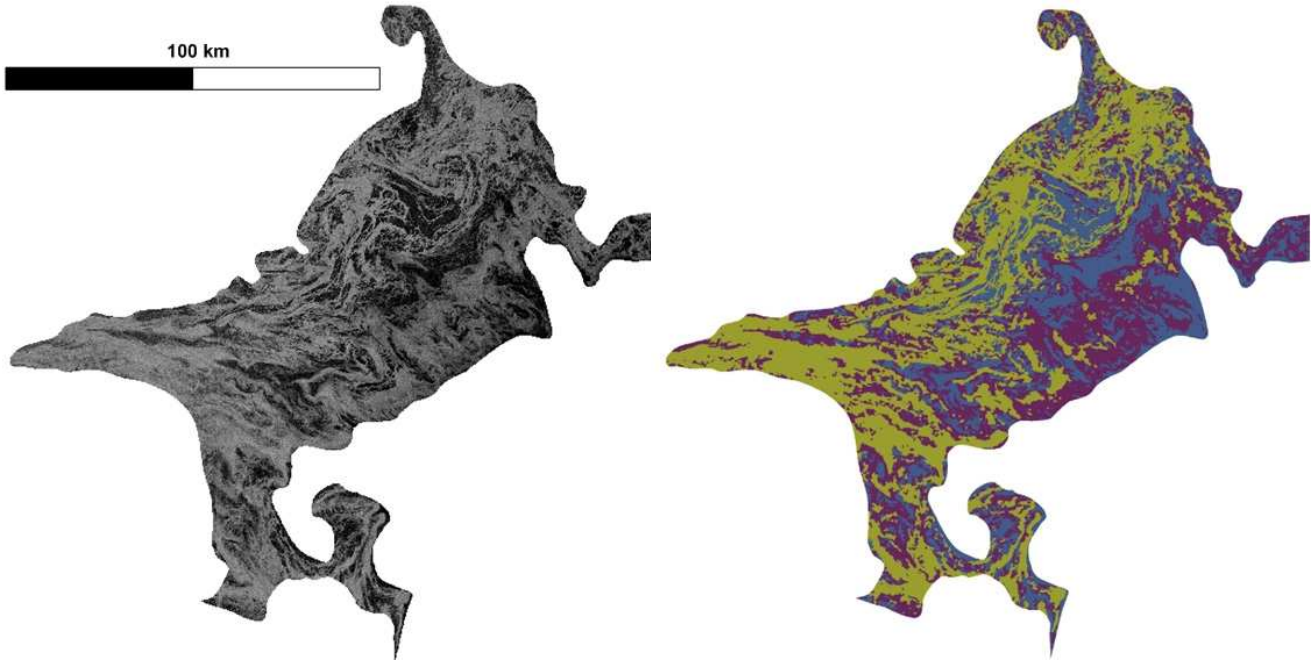
	Image 1	Image 2	Image 3
Pixel Accuracy	100%/1	93.56% /0.9140	77.50%/ 0.7012
Region Accuracy	100%/1	94.44%/ 0.9266	81.25%/ 0.7752
Accuracy Ratio	29/29	34/36	26/32

given the difficult and time consuming nature of manual segmentation. IRGS is able to successfully segment this polygon and shows similar performance for polygons summarized in Table III.

Table IV shows the performance accuracy of the three images. Image 1 has 100% accuracy,Image 2 has over 90% accuracy and Image 3 has about 80% accuracy. Even though these full SAR scenes show a high degree of intra-class variability, the classification algorithm is successful since 80% is deemed an acceptable operational rate by CIS personnel. Fig. 9 shows the original SAR (left) and classified (right) images. Note that all three images are classified automatically using exactly the same algorithm and same algorithm parameters. The full SAR sea-ice images are too large (5Kx5K) to be shown at full resolution and therefore the WMO color coded results are presented to visualize the outputs. Nevertheless, the continuity of sea-ice labels over the polygon boundaries is obvious even with such a coarse resolution. The method produced labeling continuous over polygon boundaries, using both feature and prior terms. Image 2 has a misclassification of open water with new ice in a polygon on the right side of (Fig. 9c) circled and marked with X. Intuitively, this polygon could have been perfectly labelled if there had been another polygon containing open water adjacent to it. Probably, the evaluator also differentiated the labels in that polygon based on proximity to open water. To verify, we have performed an experiment by adding a polygon with an open water adjacent to the misclassified polygon and this generates 100% accuracy. SAR data is notorious for high interand intra-class feature variability, and with such data spatial proximity is an essential criterion for labeling success. If some guidance exists for drawing the polygons the performance of the labeling algorithm increases.

Image 3 (Fig. 9f) has seven ice labels found in only 13 polygons. One polygon inside a white circle on the far left of the image in Fig. 9e is completely isolated. Some of the polygons are stretched and have shorter boundaries with each other which makes region label inferencing difficult. Such a configuration causes general misclassification of new ice with open water which is not a significant drawback from an operational perspective. The rest of the ice types: old, second-year, multi-year, grey-white, grey, have been labelled successfully in Image 3. This again stresses the importance of spatial interaction between the polygons. Overall, high classification rates have been achieved labeling the set of operational SAR sea-ice images.

Generally, if the ice analyst provided information is accurate (number of ice types, ice type labels, boundaries



(a) (b)

Fig. 8. Example of segmentation using IRGS [11]. This Image 2 polygon is approximately $2K \times 2K$ pixels and has been downsized 25 times to fit page. The polygons egg code indicates presence of grey-white, grey and new ice.(b) The polygon in (a) successfully segmented to three classes using IRGS [11]. The segmentation performance of such complex imagery demonstrates the capability of IRGS as the segmentation technique. Note that the image had to be rescaled by 13% which removes scene details.

TABLE V
COMPUTATIONAL TIME OF THE PROPOSED LABELING ALGORITHM.

	Image 1	Image 2	Image 3
Iteration θ	100	100	100
Regions n_r	29	36	32
Polygons n_p	12	13	13
Size	5028×5387	4994×5417	3818×4688
Time	$31s$	$26s$	$77s$

between polygons) then the segmentation has a stronger accuracy and the labeling process produces a more accurate pixel-level classified map.

C. Role of Prior Model

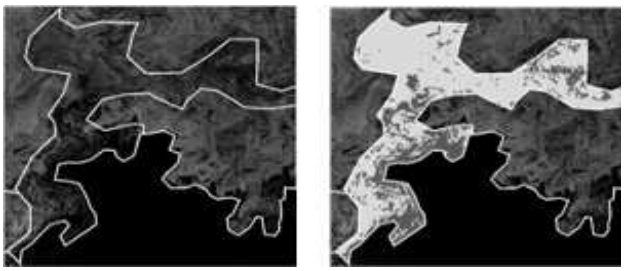
To demonstrate the role of the prior term, a comparison between using both feature and prior terms (Fig. 9a, duplicated for side-by-side viewing in Fig. 10a) and just the feature term (Fig. 10b) is performed. Here, the feature model alone is not sufficient to accurately label the regions. By removing the spatial prior edge model, only the Gaussian mixture model (GMM) remains. The labeling result using GMM alone (Fig. 10b) with accuracy 33.10% and kappa 0.1059 shows more then half of the regions have been misclassified compared to using both feature and spatial term (Fig. 10a) with accuracy/kappa 100%/1. This clearly indicates that the spatial interaction of polygons is essential in the overall model to generate accurate labeling.

D. Computational time

The overall complexity of just the labeling algorithm is $O(\theta n_r n_p + MN)$. Table V summarizes the computational time required to achieve labeling results for every test image after the segmentation. The segmentation of all polygons is 10 minutes on average. Labeling was performed using MATLAB with a 2.3 GHz Intel dual core processor, the benchmarks indicate high computational feasibility. As such, the minimal computation time supports the algorithms use in an operational environment.

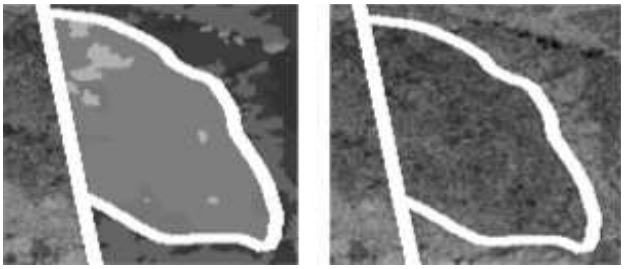
VII. CONCLUSION

A novel method has been designed and implemented for operational SAR sea-ice image classification. In the classification process, the SAR sea-ice images are segmented and labeled utilizing provided polygon data. The sea-ice classification algorithm is automatic and does not require training data but more effectively uses the joint information from all polygons to find the optimal configuration of labels based on an objective function. The objective function is defined as a combination of feature and prior models to better reflect the statistical and spatial proximity of regions.

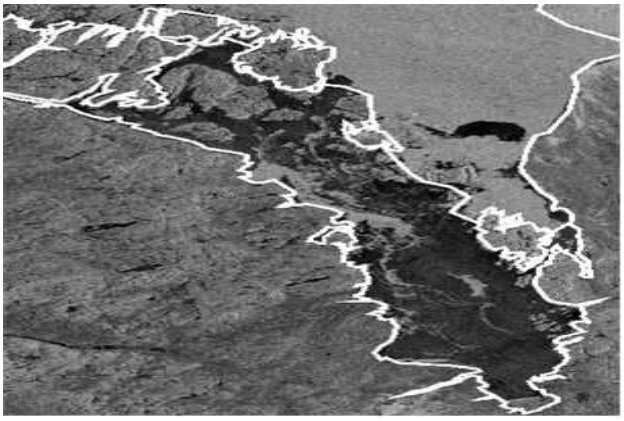


(a)

(b)



(c)



(d)

Fig. 6. Evaluator discarded polygons. (a) Image1 polygon. Requires fine grey and new ice segmentation. (b) Another Image 1 polygon. Complicated scene not visually interpretable. (c) One of three Image 3 polygons with overly smoothed segmentations.(d) One of five Image 3 polygons discarded due to land map errors.

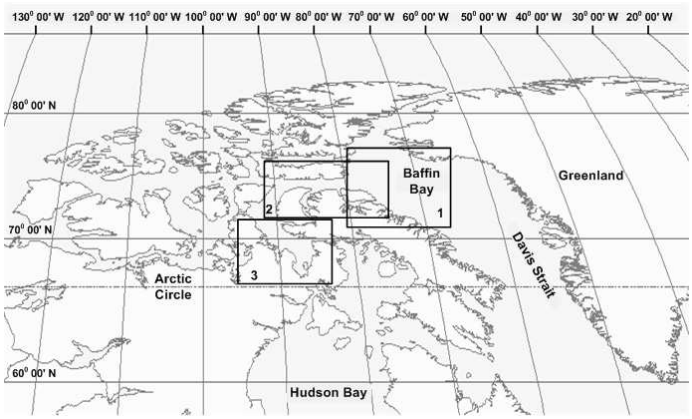


Fig. 7. Location of three reference images on map.

The concept has been demonstrated with operational SAR images provided by CIS. Using a performance evaluation framework these images have been evaluated by the trained ice expert and used as a reference to assess the sea-ice labeling performance. Consistently high performance has been obtained. The discussed framework is the only end-to-end process known to us for automatic classification of SAR sea-ice imagery.

Appendix A

DERIVATION OF REGION BASED ENERGY EQUATION

Assume the parameters u_{r_i} ,$\sigma^2_{r_i}$, k_{r_i} are known for the region r_i. Then, compute the feature model energy $E(r_i)$ for region r_i for any estimates $u_{l_{r_i}}$ and $\sigma^2_{l_{r_i}}$.

$$E(r_i) = \sum_{k=1}^{k_{r_i}} \left\{ \frac{1}{2} \ln 2\pi\sigma^2 l_{r_i} + \frac{1}{2} \frac{(x_{s_k} - u_{l_{r_i}})^2}{\sigma^2_{l_{r_i}}} \right\} \tag{17}$$

$$\begin{aligned} &= \sum_{k=1}^{k_{r_i}} \left\{ \frac{1}{2} \ln 2\pi\sigma^2 l_{r_i} + \frac{1}{2} \frac{(x_{s_k} - u_{l_{r_i}} - u_{r_i} + u_{r_i})^2}{\sigma^2_{l_{r_i}}} \right\} \\ &= k_{r_i} \frac{1}{2} \ln 2\pi\sigma^2_{l_{r_i}} + \frac{k_{r_i}}{2\sigma^2_{l_{r_i}}} (u_{r_i} - u_{l_{r_i}})^2 + \\ &\quad \frac{1}{2\sigma^2_{l_{r_i}}} \sum_{k=1}^{k_{r_i}} (x_{s_k} - u_{r_i})^2 + \frac{(u_{r_i} - u_{l_{r_i}})}{\sigma^2_{l_{r_i}}} \sum_{k=1}^{k_{r_i}} (x_{s_k} - u_{r_i}) \\ &= k_{r_i} \frac{1}{2} \ln 2\pi\sigma^2_{l_{r_i}} + \frac{k_{r_i}}{2\sigma^2_{l_{r_i}}} (u_{r_i} - u_{l_{r_i}})^2 + \\ &\quad \frac{(k_{r_i} - 1)\sigma^2_{r_i}}{2\sigma^2_{l_{r_i}}} \end{aligned} \tag{18}$$

for very large k_{r_i} , $k_{r_i} - 1 \simeq k_{r_i}$.

Acknowledgment

Acknowledgements are extended for support of this project by Geomatics for Informed Decisions (GEOIDE, a Network of Centres of Excellence under the Natural Sciences and Engineering Research Council) and CRYSYS (CRYospheric SYStem in Canada). CIS is thanked for their provision of operational SAR sea-ice images. Sea-ice expert and evaluator T. Zagon of CIS is thanked for applying the performance evaluation framework to the provided SAR sea-ice datasets and obtaining the validated data.

References

[1] "World Meteorological Organization (wmo) website," March 2009. [Online]. Available: http://www.wmo.int/

[2] D. G. Barber and E. F. Ledrew, "SAR sea ice discrimination using texture statistics: A multivariate approach," *Photogrammetric Engineering and Remote Sensing*, vol. 57, no. 4, pp. 385–395, 1991.

[3] M. Shokr, "Evaluation of second-order texture parameters for sea ice classification from radar images," *Journal of Geophysical Research*, vol. 96, 1991.

[4] J. Karvonen, M. Simila, and I. Heiler, "Ice thickness estimation using sar data and ice thickness history," in *IEEE International Geoscience and Remote Sensing Symposium 2003 (IGARSS'03), Toulouse, France*, vol. I, 2003, pp. 74–76.

[5] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society*, vol. 39, 1986.

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

[7] "Canadian ice service website [online]," March 2009. [Online]. Available: http://iceglaces.ec.gc.ca/,

[8] D. G. Flett, K. J. Wilson, P. W. Vachon, and J. F. Hopper, "Wind information for marine weather forecasting from RADARSAT-1 synthetic aperture radar data: Initial results from the marine winds from SAR demonstration project," *Canadian Journal of Remote Sensing*, vol. 28, no. 3, pp. 490–497, 2002.

[9] L. Morena, C. Pearce, L. Olfert, C.Elder, and P. Rolland, "RADARSAT-2 order handling and mission planning," *Canadian Journal of Remote Sensing*, vol. 30, no. 3, pp. 304–320, 2004.

[10] "Radarsat-1 overview website [online]," March 2009. [Online]. Available: http://ccrs.nrcan.gc.ca/

[11] 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, pp. 2126–2139, 2008.

[12] A. Bogdanov, "Neuroinspired architecture for robust classifier fusion of multisensor imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 5, pp. 1467–1487, 2008.

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

[14] D. Haverkamp, L.-K. Soh, and C. Tsatsoulis, "A dynamic local thresholding technique for sea ice classification," in *Geoscience and Remote Sensing Symposium, 1993. IGARSS '93. Better Understanding of Earth Environment., International*, vol. 2, aug 1993, pp. 638 –640.

[15] L.-K. Soh, C. Tsatsoulis, S. Member, D. Gineris, and C. Bertoia, "ARKTOS : An intelligent system for sar sea ice image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 1, pp. 229–248, 2004.

[16] F. Fetterer, C. Bertoia, and J. P. Ye, "Multi-year ice concentration from RADARSAT," in *Geoscience and Remote Sensing, 1997. IGARSS '97. Remote Sensing - A Scientific Vision for Sustainable Development., 1997 IEEE International*, vol. 1, aug 1997, pp. 402 –404 vol.1.

[17] D. A. Clausi, A. Qin, M. Chowdhury, P. Yu, and P. Maillard, "MAGIC: MAp-Guided Ice Classification System," *in Canadian Journal of Remote Sensing*, 2009.

[18] L.-K. Soh and C. Tsatsoulis, "Unsupervised segmentation of ERS and RADARSAT sea ice images using multiresolution peak detection and aggregated population equalization," *International Journal of Remote Sensing*, vol. 20, no. 15-16, pp. 3087–3109, 1999. [Online]. Available: http://www.ingentaconnect.com/content/tandf/tres/1999/00000020/F0020015/art00013

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

[20] P. Maillard, D. A. Clausi, , and H. Deng, "Map-guided sea ice segmentation and classification using SAR imagery and a MRF segmentation scheme," *," IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 12, pp. 2940–2951, 2005.

[21] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*. Boston, London: Artech House Publishers, 1998.

[22] S. Z. Li, *Markov Random Field Modeling in Computer Vision*. Springer-Verlag, 1995. [Online]. Available: http://portal.acm.org/citation.cfm?id=525898

[23] P. Fieguth, *Statistical Image Processing and Multidimensional Modeling*. Springer, 2010.

[24] M. Petrou and P. Sevilla, *Dealing with Texture*. John Wiley & Sons Inc, 2006.

[25] R. Duda, P. Hart, and D. Stork, *Pattern classification*. NY John Wiley, 2001.

[26] Q. Yu, "Automated sar sea ice interpretation," Ph.D. dissertation, Vision and Image Processing (VIP) Research Lab, University of Waterloo, Canada, 2006. [Online]. Available: http://vip.uwaterloo.ca

[27] R. Gonzales and R. Woods, *Digital Image Processing*. Prentice-Hall, 2002, no. 28.

[28] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Method*. Springer, New York, 1995.

[29] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, "Equations of state calculations by fast computing machines," *Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1091, 1953.

[30] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.

[31] J. Zhang, "The mean field theory in EM procedures for Markov random fields," *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2570–2583, 1992.

[32] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, no. 10, pp. 1–38, 1977.

[33] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. Wiley New York, 2008.

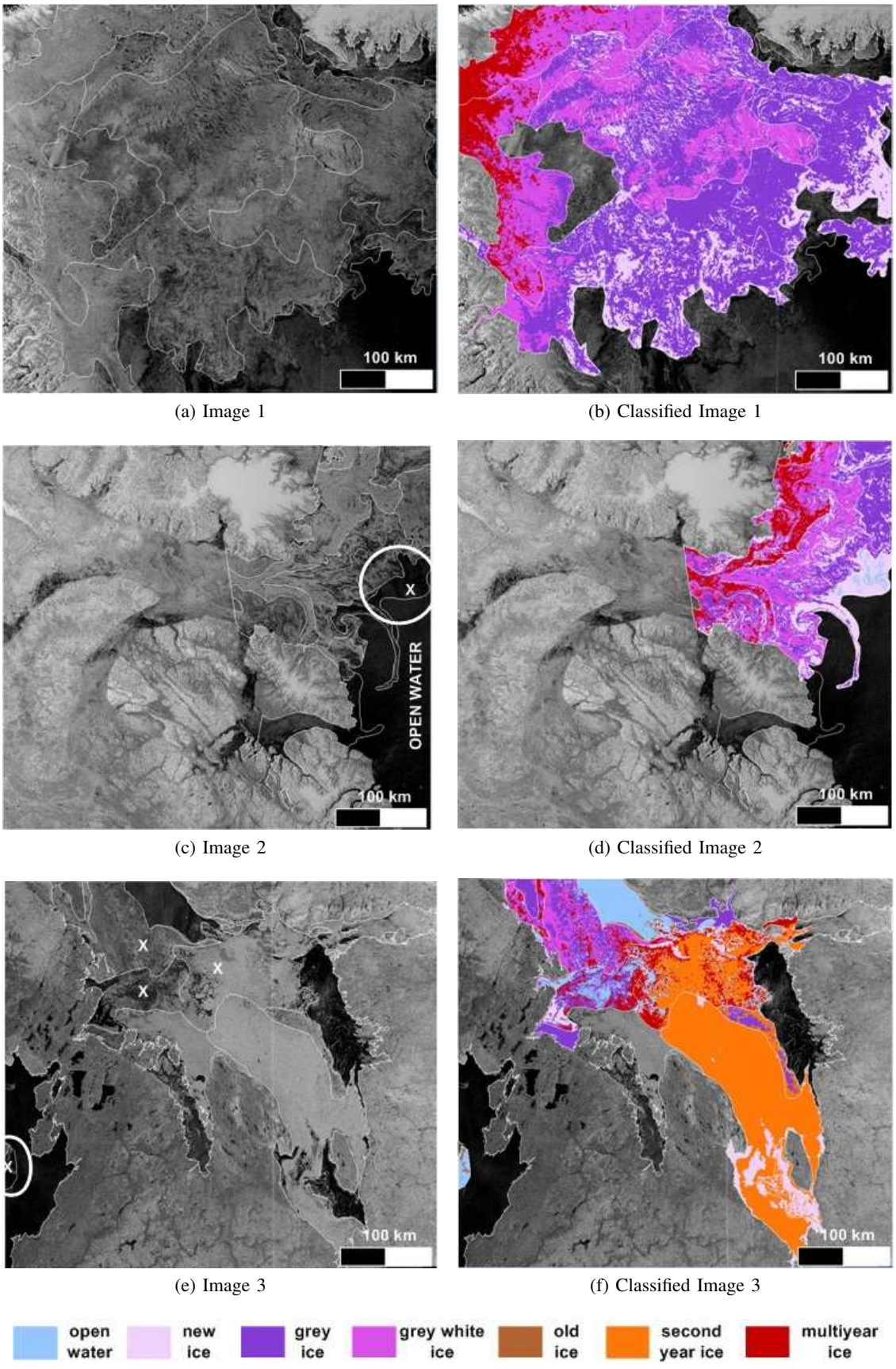
[34] H. Deng and D. A. Clausi, "Unsupervised image segmentation using a simple MRF model with a new implementation scheme," in *ICPR 2004. Proceedings of the 17th International*

Conference on Pattern Recognition, vol. 2, aug. 2004, pp. 691 – 694 Vol.2.

[35] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and psychological measurement*, vol. 20, pp. 37–46, 1960.

Shuhratchon Ochilov Shuhratchon Ochilov graduated from Istanbul Technical University with a B.Sc. in Electronics and Telecommunication Engineering in 2003. He has been appointed as a research assistant from 2004-2006 at the University of South Alabama where he has received M.Sc. in Electrical and Computer Engineering. During his graduate studies he has been nominated by Phi Beta Delta Honor Society as an International Scholar at the University of South Alabama. Currently, he has been pursuing doctoral studies at the University of Waterloo, Systems Design Engineering. During his studies he has received numerous awards and scholarships. His area of interest is in remote sensing, automated classification, digital signal/image processing, automatic target recognition/tracking, machine learning and multiresolution analysis.

David A. Clausi David A. Clausi (S93-M96-SM03) received the B.A.Sc. M.A.Sc. and Ph.D. degrees in systems design engineering from the University of Waterloo, Waterloo, ON, Canada, in 1990, 1992, and 1996, respectively. After completing his doctorate, he worked in the medical imaging field at Agfa (Waterloo). He started his academic career in 1997 as an Assistant Professor in Geomatics Engineering at the University of Calgary, Alberta, Canada. In 1999, he returned to his alma mater and is now a Professor specializing in the fields of Intelligent and Environmental Systems. He is an active interdisciplinary and multidisciplinary researcher. He has an extensive publication record, publishing refereed journal and conference papers in the diverse fields of remote sensing, computer vision, algorithm design, and biomechanics. His primary research interest is the automated interpretation of synthetic aperture radar (SAR) sea-ice imagery, in support of operational activities. His research efforts have led to successful commercial implementations. Dr. Clausi was the co-chair of IAPR Technical Committee 7-Remote Sensing during 2004-2006. He has received numerous scholarships, paper awards, and two Teaching Excellence Awards



(a) Image 1

(b) Classified Image 1

(c) Image 2

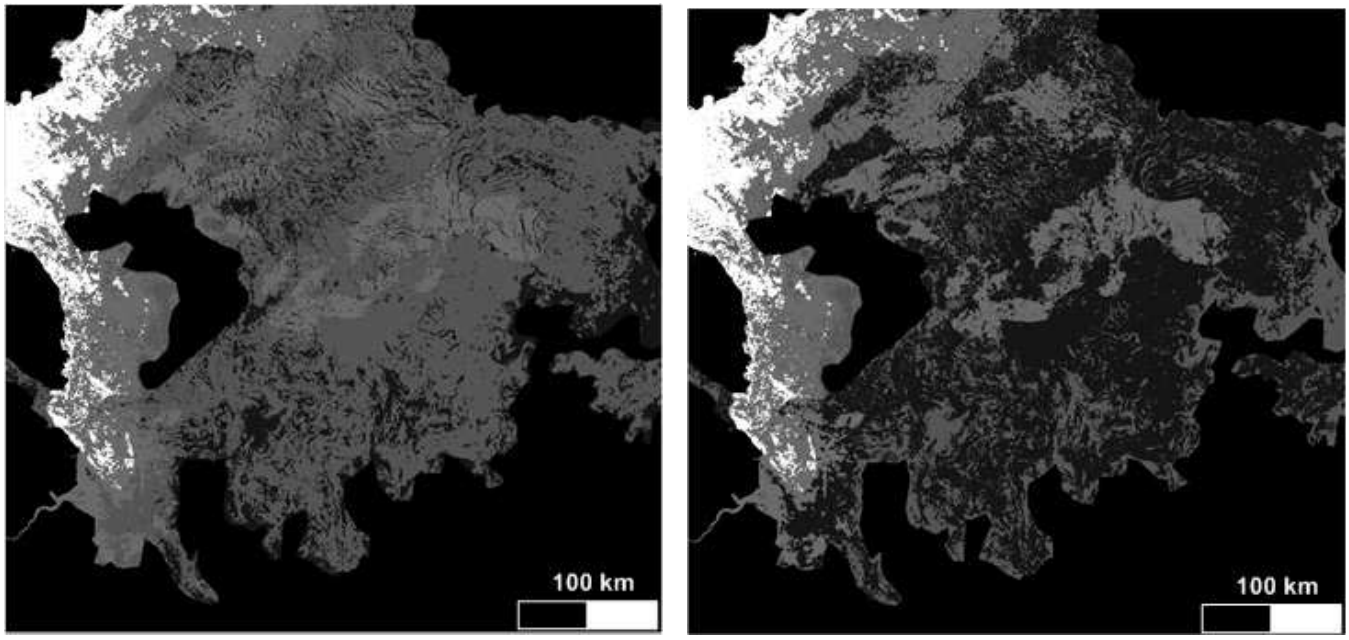
(d) Classified Image 2

(e) Image 3

(f) Classified Image 3

(g) WMO [1] color code

Fig. 9. Classification of SAR sea-ice images with proposed technique. Polygon boundaries are overlaid on the image as white contours. Images are too large($5K \times 5K$) to show details but this segmentation and labeling success can be observed.



(a) (b)

Fig. 10. The effect of spatial term on classification performance. (a) The labeled output using the proposed algorithm with accuracy/kappa 100%//1. (b) The labeled output using just the GMM feature model with accuracy 33.10% and kappa 0:1059. Without the spatial context model, the labeling process produces poor results.