A Semi-supervised Approach for Ice-water Classification Using Dual-Polarization SAR Satellite Imagery

Fan Li, David A. Clausi, Lei Wang, and Linlin Xu
University of Waterloo
200 University Ave W, Waterloo, ON N2L 3G1
{f33li, dclausi, lei.wang, l44xu}@uwaterloo.ca

Abstract

The daily interpretation of SAR sea ice imagery is very important for ship navigation and climate monitoring. Currently, the interpretation is still performed manually by ice analysts due to the complexity of data and the difficulty of creating fine-level ground truth. To overcome these problems, a semi-supervised approach for ice-water classification based on self-training is presented. The proposed algorithm integrates the spatial context model, region merging, and the self-training technique into a single framework. The backscatter intensity, texture, and edge strength features are incorporated in a CRF model using multi-modality Gaussian model as its unary classifier. Region merging is used to build a hierarchical data-adaptive structure to make the inference more efficient. Self-training is concatenated with region merging, so that the spatial location information of the original training samples can be used. Our algorithm has been tested on a large-scale RADARSAT-2 dual-polarization dataset over the Beaufort and Chukchi sea, and the classification results are significantly better than the supervised methods without self-training.

1. Introduction

The operational mapping of sea ice is beneficial for several important purposes, including ship navigation, weather forecasting, and environmental science. Among many satellite sensors, synthetic aperture radar (SAR) is very suitable for sea ice mapping because of its all-weather and all-day imaging capability. Also, the recently-launched RADARSAT-2 satellite makes dual-band polarization on large scale available, so that more information can be provided for the operational sea ice mapping.

In the Canadian Ice Service (CIS), the interpretation of sea ice imagery is currently performed manually by ice analysts everyday. An example of the manual ice chart is shown in Fig. 1. The manually-outlined irregular polygons are "egg codes" defined by World Meteorological Organization, which record the ice conditions such as the ice concentration in the polygons estimated by the ice experts based on their experiences. The ice chart provides good interpretation in a large scale, but is lack of details because the ice-water boundaries are coarsely outlined. Also, the result may not be consistent for different ice analysts [14]. By comparison, automated interpretation approaches have the potential of generating high-volume pixel-level classification maps with no inter-operator bias.

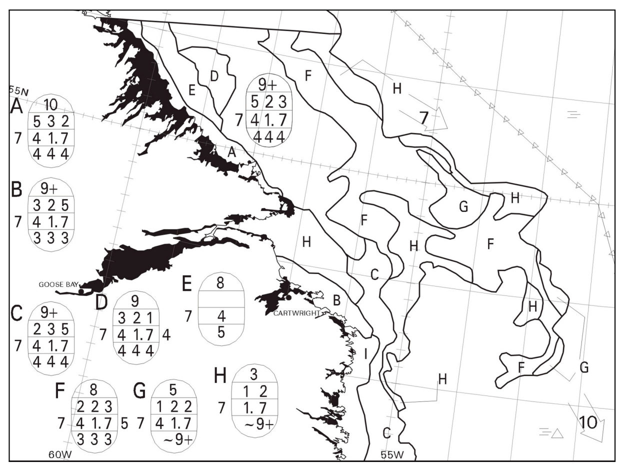


Figure 1. Example of an ice chart [8] and the egg codes, recording the total ice concentration, the partial ice concentration, the stage of ice development, and the form of ice in the polygon.

The automated classification of SAR sea ice imagery has been studied for over a decade [6, 10–12, 18, 21]. However, there is no existing algorithm that can be in operational use so far. There are mainly two reasons. On one hand, the properties of ice and open water in the SAR imagery are affected by many factors such as incidence angle, snow condition, and wind speed, resulting in tremendous within-class variability across scenes and even within a scene. Also, SAR imagery is corrupted by speckle noise, which degrades the image details and further decreases class separability. As a result, traditional pixelwise classification methods [10, 11] would be incapable of achieving satisfactory result. On the other hand, traditional supervised classification methods require sufficient labeled samples for training. However, the ground-truthing requires the expertise and experience of interpreting sea ice, and it is a very time-consuming and tedious task considering the large size of SAR images. Otherwise, if only a small amount of pixels are labeled for training, they might not enough for characterizing the true data distribution and thus yield to inferior classification results.

In recent years, semi-supervised learning techniques have been applied to remote sensing image classification tasks [2–4, 13]. These methods incorporate additional unlabeled pixels into training to compensate the insufficiency of the original training samples. However, these methods usually require high computational cost, and the improvement is sometimes limited if "bad" unlabeled samples are selected for training. Recently, Dópido et al. [7] used a self-training technique for classification of hyperspectral imagery, and only pixels that are near the initial training samples in their spatial coordinates can be selected for training. In the remote sensing context, the selection of training samples and the classification are usually performed in the same image, and therefore the spatial location information of the original training samples can be used to guide the selection of new training samples. According to the Tobler's first law of geography [19], pixels that are near the original training samples are more likely to be in the same class as the training samples, and are thus more reliable if they are used for self-training.

Motivated by [7], a semi-supervised approach called self-training IRGS (ST-IRGS) is presented in this paper. It inherits some useful properties of the iterative region growing using semantics (IRGS) algorithm, an unsupervised segmentation algorithm that has been successfully applied for SAR sea ice imagery [21]. In the ST-IRGS, hierarchical region merging is integrated with a conditional random field (CRF) to iteratively reduce the number of nodes, and edge strength is used in both classification and region merging. The key feature of the ST-IRGS is an embedded self-training procedure. Compared to [7], the ST-IRGS can iteratively expand the training candidate set owing to its region merging property, so that the abundant unlabeled samples can be explored even if they are not near the original training samples. Also, the correctness of the labels can be ensured by the region properties.

The rest of the paper is structured as follows. Section 2 describes the proposed ST-IRGS algorithm in the context of the ice-water classification problem. Experiments on a RADARSAT-2 SAR dual-polarization dataset are reported in Section 3. Section 4 concludes and outlines a direction for future work.

2. Methodologies

2.1. Problem formulation

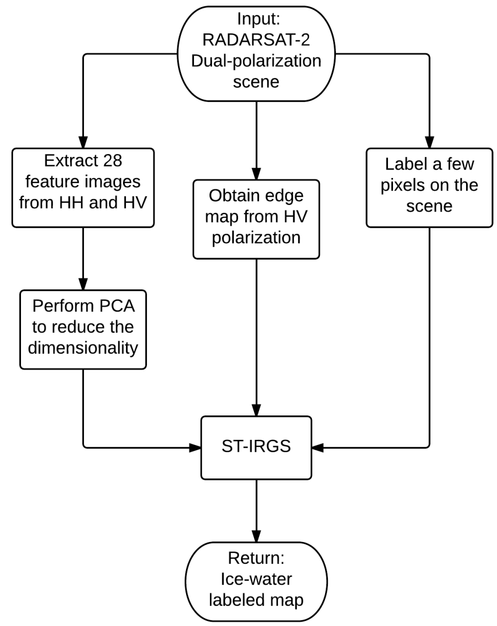


Figure 2. Overview of the ice-water classification using the ST-IRGS algorithm

Fig. 2 shows an overview of the ice-water classification using the proposed ST-IRGS algorithm. Given a few labeled pixels in an image, we aim to learn the optimal labeling configuration of all the pixels in the image. For a first-order CRF model, the posterior probabilities $P(\mathbf{y} \mid \mathbf{x})$ can be formulated as [9, 17]:

$$\begin{aligned}\log P(\mathbf{y} \mid \mathbf{x}, \theta) &= \sum_{\mathbf{s}} \phi_{\mathbf{s}}(\mathbf{y_s}, \mathbf{x}^{\phi}; \theta^{\phi}) \\ &+ \beta \sum_{s} \sum_{t \in \eta_s} \xi_{st}(y_s, y_t, g_{st}(x^{\xi}); \theta^{\xi}) - \log \mathbf{Z}(\theta, \mathbf{x})\end{aligned} \tag{1}$$

where $\phi_s(\cdot)$ and $\xi_{st}(\cdot)$ are unary and pairwise clique potentials respectively, s indexes nodes in a discrete rectangular lattice, η_s refers to the 4-connected neighbors of node s, $g_{st}(x^{\xi})$ is the edge feature for two adjacent nodes s and t, $\mathbf{Z}(\theta, \mathbf{x})$ is a partition function, $\{\mathbf{x}^{\phi}, \mathbf{x}^{\xi}\} \subset \mathbf{x}$ are data for the potentials, β is a weight parameter, and $\theta = \left\{\theta^{\phi}, \theta^{\xi}, \beta\right\}$ are model parameters.

In our approach, we use a combination of backscatter intensity and texture features for $\mathbf{x}^{\phi}$ in the unary potentials. In the previous literature, grey-level co-occurrence matrix (GLCM) parameters have been demonstrated to be effective in distinguishing different ice types and open water [5]. In our approach, we adopt a total of 28 features including the mean, standard deviation, and GLCM measures in different window sizes extracted for ice-water classification using a RADARSAT-2 image dataset [12].

The unary potentials are defined using Gaussian models. Even if the distribution of features is not strictly Gaussian, the Gaussian models can still be used to approximate it [6, 21]. For SAR sea ice imagery, a single Gaussian modality is usually insufficient of modeling a class even within a scene. Fig. 3 shows an example that the principal component of the water class has multiple mixtures. Therefore, we use multiple Gaussian mixtures to model each class. The unary potentials are defined as:

$$\begin{aligned}&\phi_s(y_s, \mathbf{x}^{\phi}; \alpha_{\mathbf{k}}, \mu_{\mathbf{k}}, \mathbf{\Sigma_k}, \mathcal{C}_{\mathbf{y_s}})\\ &= \log\left\{p(y_s)\sum_{k}^{\mathcal{C}_{y_s}} \alpha_k \mathbb{N}(\mathbf{x}_\mathbf{s}^{\phi}, \mu_{\mathbf{k}}, \mathbf{\Sigma_k})\right\}\end{aligned} \tag{2}$$

where $p(y_s)$ is a class prior, $\mathcal{C}_{y_s}$ is the number of mixtures in the class, $\mu_{\mathbf{k}}$ and $\mathbf{\Sigma_k}$ are mean and variance of the mixture k, and α_k is the mixture prior in the class $\mathcal{C}_{y_s}$. Although other state-of-the-art classifiers can be used as the unary classifier, the multi-modality Gaussian model performs very well for ice-water classification in our experiment given sufficient training samples, and the estimated mixture parameters can be used for region merging and determining the weight parameter later.

For x_ξ in the pairwise potentials, we only use the backscatter intensity in the HV polarization which is less sensitive to incidence angle effect compared to the HH polarization. The edge feature is set to measure the gradient between neighboring pixels [20]. Thus, the pairwise potential is defined as:

$$\xi_{st}(y_s, y_t, \mathbf{x}) = \begin{cases}\beta g_{st}(x^{\xi}) & y_s \neq y_t\\ 0 & \text{otherwise}\end{cases} \tag{3}$$

$$g_{st}(x^{\xi}) = exp\left[-\left(\frac{x_s^{\xi} - x_t^{\xi}}{K}\right)^2\right] \tag{4}$$

where β is a weight parameter, and K is a gradually increasing parameter during the optimization iterations, which has been demonstrated to be experimentally satisfactory for SAR sea ice imagery [20].

2.2. Problem solving

An alternating procedure is used to perform both parameter estimation and inference in (1). Fig. 4 shows the flow chart of the ST-IRGS algorithm. In each iteration, the model parameters are first fixed, and the label configuration is optimized. Then, the model parameters are updated based on the current labels. In the parameter estimation step, the parameters related to the unary potentials are estimated independently. The number of mixtures is estimated using Bayesian information criterion [16], and then α_k, $\mu_{\mathbf{k}}$, and $\mathbf{\Sigma_k}$ are estimated using expectation maximization (EM). After the parameters are estimated, Gibbs sampling is used for inference. Each node is processed only once in one iteration.

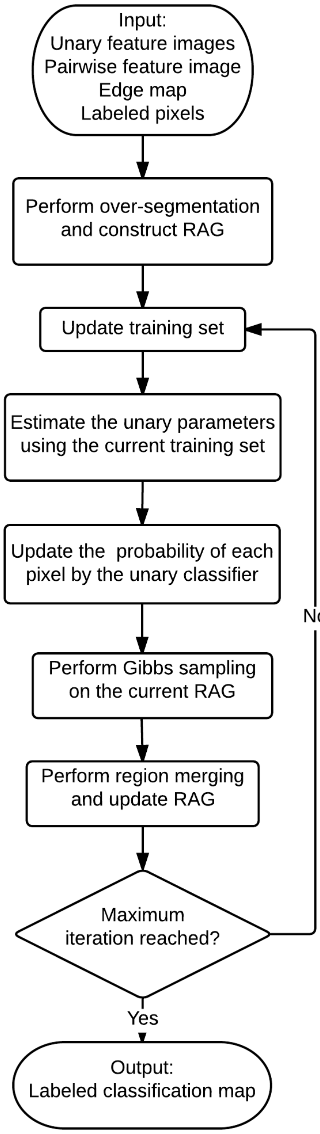
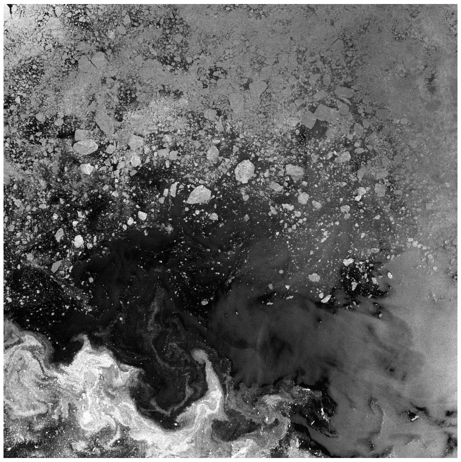
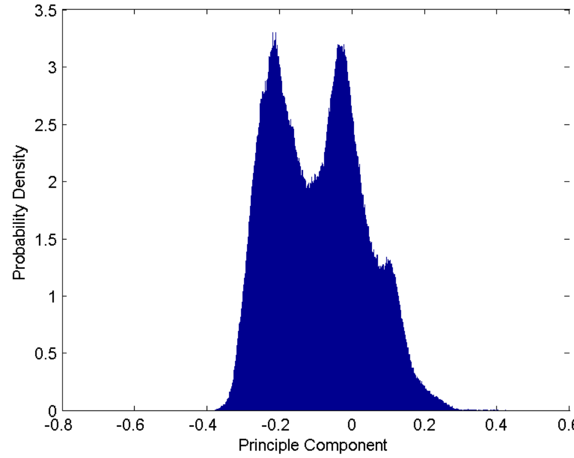


Figure 4. Flow chart of the ST-IRGS algorithm

Similar to the IRGS algorithm [20], we incorporate the hierarchical region growing procedure into the iterations to build a hierarchical data-adaptive structure in order to make the optimization more efficient. A region adjacency graph (RAG) is first generated from the initial over-segmentation



(a)



(b)

Figure 3. (a) RADARSAT-2 SAR imagery in HH polarization captured over Beaufort sea on July 30, 2010. (b) Probability density function of the principal component of the 28 features for the water class, which clearly shows multiple mixtures due to the incidence angle effect.

result. Then, similar neighboring pixels are iteratively merged into regions, so that the number of nodes can be significantly reduced, and thus the convergence rate can be increased. In our approach we also assume each region to be single-modality Gaussian distribution, so the same merging criterion as the IRGS can be used:

$$\partial E_{ij} = \sum_{s\in\Omega_k} \log(\sigma_k) - \sum_{s\in\Omega_i} \log(\sigma_i) - \sum_{s\in\Omega_j} \log(\sigma_j) \\ -\beta \sum_{\langle s,t\rangle \in C; s\in\Omega_i, t\in\Omega_j} g_{st}(x^{\xi}) \tag{5}$$

where Ω_i and Ω_j are two neighboring regions for merging, Ω_k is the region after merged, and σ is the variance of the region.

The detailed derivation of (5) can be found in [20]. To preserve the single-modality property of the regions, each region is assigned a mixture label based on the previous estimated number of mixtures and Gaussian parameters. Only regions that are in both the same predicted class label and mixture label are allowed to be merged. In each iteration, the edges of adjacent nodes with negative ∂E_{ij} are put into a merging list in an ascending order to make the merging more efficient, and the corresponding region and edge information is updated after each merging. Also, each node is only allowed to be merged once in the first a few iterations to avoid extremely-large regions that may result in imbalanced number of training samples. For SAR sea ice imagery and other kinds of remote sensing imagery that contain homogeneous regions, the number of nodes can be finally reduced to a very small number, so that even a very simple optimization method can achieve near-optimal solutions.

To address the problem of insufficient training samples, a self-training procedure is followed by region merging in each iteration. Self-training is a kind of semi-supervised technique that iteratively retrains the classifier using the predictions that are confident [22]. However, traditional self-training methods are largely dependent on the performance of the classifier. If the classifier performs very badly, the predictions are not be reliable and will degrade the classification performance if they are used for retraining. In our approach, we only select pixels which are in the same regions as the original training samples to be new training samples.

Another important issue is to determine the weight parameter β between the unary and pairwise potentials. Traditional classification approaches usually seek for the optimum weight parameter using grid search, which is time-consuming and unreliable when only limited training samples are available. In the ST-IRGS, we use the scheme for adapting the weight parameter of the IRGS. The parameter β is updated based on class separability in each iteration:

$$\beta = C_1 \frac{J/C_2}{1 + J/C_2} \beta_0 \tag{6}$$

where β_0 is the ratio of the current total class boundary length over the image size [15], C_1 and C_2 are constants, and J is the separability of ice and water classes.

To calculate the pairwise class separability, we still use the Fisher's criterion, but we need to sum up the values of all the pairwise mixtures because each class has multiple mixtures:

$$J = \sum_{i\in\mathcal{C}_{\text{ice}}} \sum_{j\in\mathcal{C}_{\text{water}}} \left\{ \alpha_i \alpha_j \frac{|\mu_i - \mu_j|}{\sigma_j^2 + \sigma_j^2} \right\} \tag{7}$$

where μ_i and σ_i are the mean and variance of the HV backscatter intensity for pixels that belong to the mixture i.

In each iteration, the β and K in (3) and (4) are updated correspondingly until a maximum number of iterations $\tau_{\max}$ is reached.

3. Experiments and Analysis

3.1. Data

The data for testing include eight scenes obtained from the C-band RADARSAT-2 SAR satellite over the Beaufort and Chukchi Sea area from May to December in the year 2010. They were captured in the ScanSAR Wide mode, which is the most useful beam mode for sea-ice monitoring. HH and HV dual-polarizations are provided in the ScanSAR Wide mode. The spatial resolution of images is 50 m, and the image sizes are around 10 000 $\times$ 10 000 pixels. The test images were acquired from both ascending and descending passes, with an incidence angle ranging from 20° to 49°. Each scene contains both ice and open water.

In the pre-processing step, the log-transformed original images were down-sampled using 4 $\times$ 4 block averaging to meet operational requirement [12]. The test images are cropped from the blocked-averaged images to eliminate the land area, and each test image has the same size of 2000 $\times$ 2000 pixels. Even though the down-sampled images yield to coarser classification results, the results are still far more detailed than the expected human interpretation [12]. The vector-based ground truth for the test images was made by an experienced ice analyst. To make a better evaluation of the algorithms, accurate pixelwise ground truth for each image has been created based on the vector-based ground truth for both training and validation.

3.2. Experimental setup

The ST-IRGS algorithm is implemented in Microsoft Visual C++ 2010. C_1 and C_2 in (6) are set to 3 and 0.4 separately. We use the same weight parameter for all the test images considering its adaptability, even though the accuracy could be improved by carefully tuning the parameters for each scene. $\tau_{\max}$ is set to 100. The setting of K in (4) is the same as the IRGS algorithm [20]. The maximum number of mixtures $\mathcal{C}_i$ for each class i is set to 5. When the number of training samples in a class is less than 200, a single-modality Gaussian model is used in order to guarantee sufficient samples for estimating the model parameters. To reduce the computational cost, we randomly select 5000 samples for parameter estimation in the EM algorithm if the number of available training samples in a class is greater than 5000, and empirically the selection of more training samples does not show additional improvement. Also, the parameters are only updated when the available training samples for a class are increased by 2%. In the first 30 iterations, each region is only allowed to be merged once.

In the experiment, only ten pixels for each class are randomly selected for training samples. We use the multi-modality Gaussian maximum likelihood classier (MGMLC) and GMRF which are both closely related to the proposed ST-IRGS algorithm for comparison. MGMLC uses the pixels in the same watershed algorithm as the original training samples for parameter estimation, and it serves as the unary classifier for both GMRF and ST-IRGS. GMRF combines MGMLC with the standard MRF model, and the graph-cut algorithm [1] is used for inferencing the labels. The weight parameter of GMRF adopts the one with highest test accuracy in a set $\{2^0, 2^1, 2^2, ..., 2^8\}$. For all the methods, PCA is first applied to the 28 features and the first five principal components are remained.

3.3. Experimental results and analysis

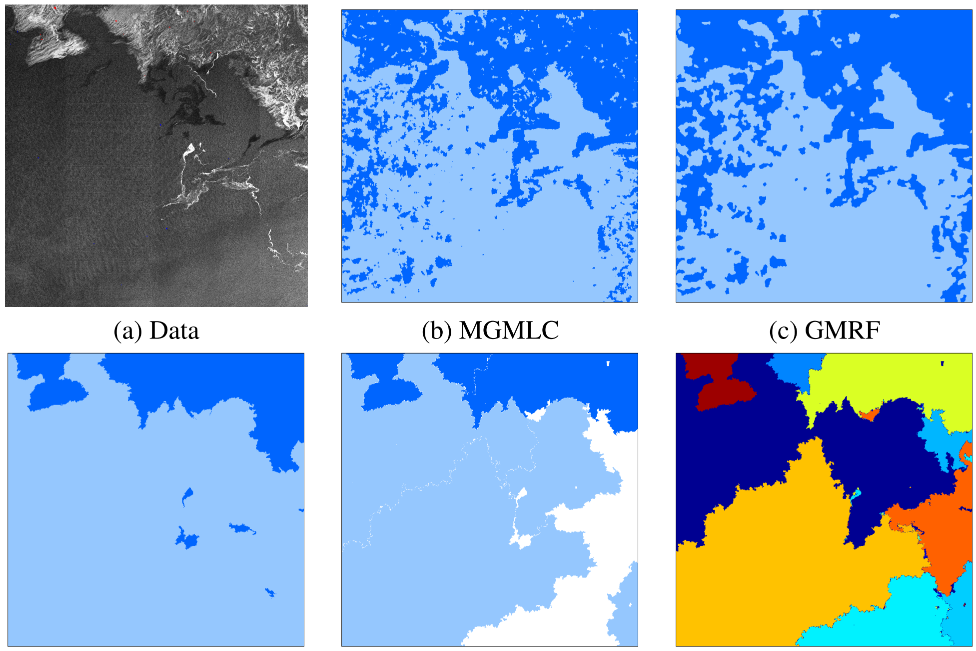
The classification result of the whole test dataset is shown in Table 1. Each image is tested for 10 times, and the averaged mean and standard deviation of the overall classification accuracy (OA) and the Kappa coefficient for all the images in the dataset are reported in the table. The ST-IRGS achieves about 95 percent classification accuracy, which is significantly higher than the other two methods. Moreover, the variation of the OA due to random sampling is significantly reduced by the ST-IRGS. This is because as the training set is expanded, the classification performance becomes less sensitive to the initial training samples. Considering that the initial samples of the ST-IRGS are also selected totally by random, the OA could be even higher in practice because an experienced human interpreter may be able to label more representative samples.

Table 1. Overall classification accuracy (OA%) and Kappa coefficient for MGMLC, GMRF, and ST-IRGS on a RADARSAT-2 SAR dual-polarization dataset.

	OA(%)	Kappa
MGMLC	78.67±9.31	0.48±0.13
GMRF	82.06±9.40	0.54±0.17
ST-IRGS	94.38±0.89	0.84±0.03

Fig. 5 shows the result on a RADARSAT-2 image scene captured over the Beaufort sea on October 3, 2010. The ice usually starts to freeze in the Beaufort Sea in early October. The air temperature was -1.0°C at the moment. In Fig. 5 (a), different stages of ice growth including new ice, grey ice, and grey-white ice are observed. There is also some multi-year ice in the North East. It had survived the summer and began to build up together with the first-year ice that is freezing. The wind speed at PRDA2 tower was 12.6 m/s, which results in the wind-roughened texture of the open water on the image.

Due to the high wind and the incidence angle effect, some areas of the open water have similar properties to ice in both intensity and texture. In Fig. 5 (b), the rough water area in the West and the open water in the near range in



(a) Data (b) MGMLC (c) GMRF

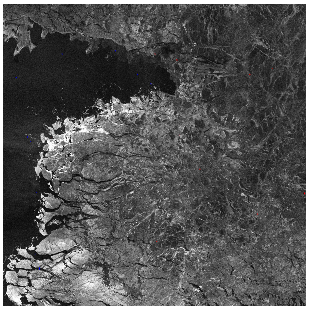
(d) ST-IRGS (e) Self-training set (f) Region map

Figure 5. Classification results of the data captured on October 3, 2010. (a) shows the HV polarization of the data, with original training samples overlain on the image (red: ice; blue: open water). (b), (c), and (d) are classification results of MGMLC, GMRF, and ST-IRGS respectively. (e) is the final training sample map expanded by self-training of ST-IRGS. (f) is the final region map of ST-IRGS.

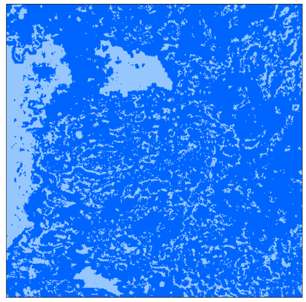
the South East are misclassified as ice. Conversely, some newly-formed ice that has similar backscatter to water is also misclassified. Moreover, class boundaries are not accurate because the edges of the texture images are blurred due to the large GLCM window size used. This is unavoidable because if smaller texture window size is used, some texture patterns might fail to be extracted. In Fig. 5 (c), GMRF can refine some of the labeling with the aid of the spatial context, but in some open water area, there is even more misclassification due to the poor estimation of the unary classifier. Also, there is no improvement on the correction of the class boundaries. In Fig. 5 (d), the classification result is significantly improved owning to the expanded training set by the self-training technique. Fig. 5 (e) shows that most pixels of the image expanded into the training set are correctly labeled. There are a few unexpanded regions in the South East due to the strong incidence angle effect, but they can be finally labeled correctly. An average of 97.2% overall classification accuracy can be achieved in 10 tests using different randomly-selected samples. Also, the class boundaries are corrected after the incorporation of edge strength into the CRF energy function.

Beyond the improvement, there are two obvious errors in Fig. 5 (d). First, the grease ice in the middle of the image is mostly misclassified. Even though a human interpreter is able to identify some grease ice by its difference from the surrounding open water, the feature space of grease ice is overlapped with that of open water in other areas of the scene. The correct labeling of the grease ice may be at the cost of misclassifying some open water, as shown in Fig. 5 (b) and (c). Also, the small pieces of ice in the South East are labeled as open water. Due to the texture window size, there is little difference between those small ice floes and the open water in the texture feature space. Also, the small regions tend to be considered as noise in a spatial context model. Nevertheless, neither of these misclassifications incurs significant operational issues [12].

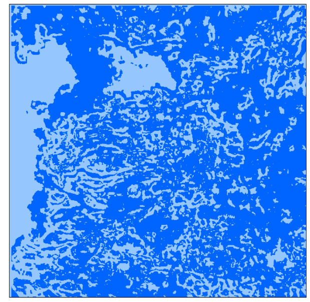
Another result is shown in Fig. 6, using the scene captured over the Chukchi sea on November 14, 2010. The air temperature was -15.1°C. The wind speed measured at PRDA2 tower was 6.7 m/s. In mid November, the ice coverage is increasing towards the South West in the Chukchi Sea, and new ice starts to form. In this image, there is a mixture of first-year ice, grey ice, and grey-white ice in the North East. Some ice area has very dark intensity, and is difficult to be distinguished from open water from in a small scale. As a result, MGMLC that is only based on small watershed regions is unable to achieve satisfactory classification result. GMRF can correct some misclassifications of open water into ice, but does not help improve the classification accuracy of ice. The ST-IRGS correctly classifies most of the pixels by expanding the training set, and achieves an OA of 96.0% over 10 tests. In Fig. 6 (e), we can see that the training set of ice is expanded to most of the ice region



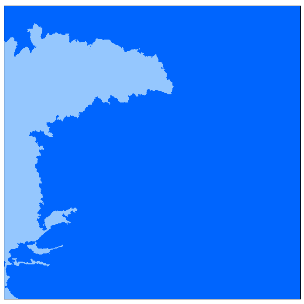
(a) Data



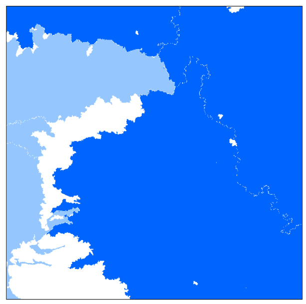
(b) MGMLC



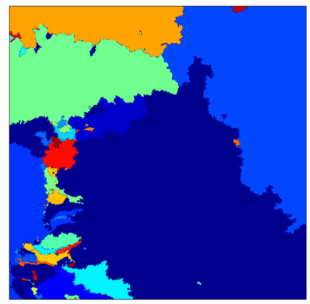
(c) GMRF



(d) ST-IRGS



(e) Self-training set



(f) Region map

Figure 6. Classification results of the data captured over Chukchi sea on November 14, 2010. (a) shows the HV polarization of the data, with original training samples overlain on the image (red: ice; blue: open water). (b), (c), and (d) are classification results of MGMLC, GMRF, and ST-IRGS respectively. (e) is the final training sample map expanded by self-training of ST-IRGS. (f) is the final region map of ST-IRGS.

except the area between ice and open water, where the ice is relatively new and the characteristics is different from the rest of ice. This area can also be correctly labeled in the final result.

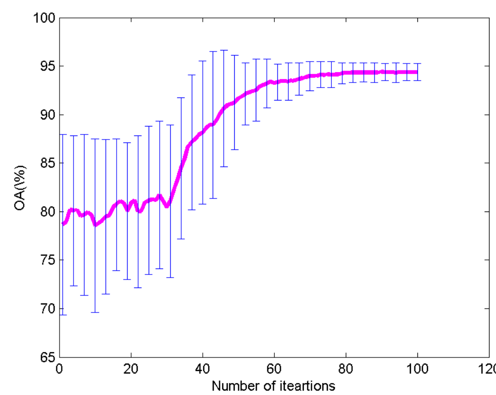


Figure 7. Overall classification accuracy by ST-IRGS during the iterations.

Fig. 7 shows the change of the OA during the iterations for the dataset. The OA fluctuates at the beginning of the iterations because even though the number of training samples increases, they are still incapable of characterizing the whole image. Once new samples are incorporated, the classifier will change greatly. Such an intermediate result is not reliable. If the self-training is only based on this result without the region constraint, the subsequent result might be even worse. Instead, the ST-IRGS only trusts the predictions in the regions that contain the original training samples, in order to guarantee the correctness of the self-training samples. After 30 iterations, the OA starts to increase as more samples are added into the training set, and become stable both after 80 iterations.

In Fig. 8, the number of regions is gradually reduced from about 100,000 at the beginning to less than 400 at the end of the iterations. There are two advantages. On one hand, the reduction of regions makes the optimization more efficient and can help for the extraction of high-level features. On the other hand, it is very convenient to correct the classification results manually by re-labeling the misclassified regions using the region maps in any previous iteration.

4. Conclusions

The ST-IRGS algorithm is capable of accurately distinguishing ice and open water in large-scale dual-polarization SAR images using a very small number of labeled samples. The multi-modality Gaussian model is suitable for describing the class distributions considering the complexity of the SAR sea ice data. The inherent combination of self-training in the IRGS framework can iteratively and correctly expand the training set and improve the classifier. Robust classification results have been achieved in testing of a RADARSAT-

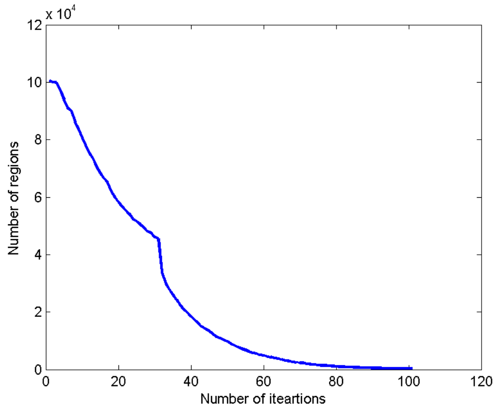


Figure 8. Number of regions by ST-IRGS during the iterations.

2 dual-polarization dataset captured over the Beaufort and Chukchi sea area. Future work includes improving the identification of grease ice and small ice floes.

5. Acknowledgement

This work is partly funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Space Agency (CSA), and the Canada Research Chairs program. The authors would like to thank MacDonald, Dettwiler and Associates for the RADARSAT-2 dataset and ice analyst D. Isaacs for providing the corresponding ground truth.

References

[1] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239, 2001.

[2] L. Bruzzone, M. Chi, and M. Marconcini. A novel transductive SVM for semisupervised classification of remote-sensing images. *IEEE Trans. Geosci. Remote Sens.*, 44(11):3363–3373, 2006.

[3] G. Camps-Valls, T. Bandos Marsheva, and D. Zhou. Semi-supervised graph-based hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.*, 45(10):3044–3054, 2007.

[4] M. Chi and L. Bruzzone. Semisupervised classification of hyperspectral images by SVMs optimized in the primal. *IEEE Trans. Geosci. Remote Sens.*, 45(6):1870–1880, 2007.

[5] D. A. Clausi and B. Yue. Comparing cooccurrence probabilities and Markov random fields for texture analysis of SAR sea ice imagery. *IEEE Trans. Geosci. Remote Sens.*, 42(1):215–228, 2004.

[6] H. Deng and D. A. Clausi. Unsupervised segmentation of synttic aperture radar sea ice imagery using a novel Markov random field model. *IEEE Trans. Geosci. Remote Sens.*, 43(3):528–538, 2005.

[7] I. Dópido, J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias, and J. A. Benediktsson. Semi-supervised self-learning for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.*, 51(7):4032–4044, 2013.

[8] D. Fequest. *MANICE: manual of standard procedures for observing and reporting ice conditions*. Environment Canada, 2005.

[9] X. He, R. S. Zemel, and M. Carreira-Perpindn. Multiscale conditional random fields for image labeling. In *CVPR*, volume 2, pages II–695, 2004.

[10] J. A. Karvonen. Baltic sea ice sar segmentation and classification using modified pulse-coupled neural networks. *IEEE Trans. Geosci. Remote Sens.*, 42(7):1566–1574, 2004.

[11] R. Kwok, E. Rignot, B. Holt, and R. Onstott. Identification of sea ice types in spaceborne synthetic aperture radar data. *Journal of Geophysical Research: Oceans (1978–2012)*, 97(C2):2391–2402, 1992.

[12] S. Leigh, Z. Wang, and D. A. Clausi. Automated ice-water classification using dual polarization sar satellite imagery. *IEEE Trans. Geosci. Remote Sens.*, 52(9):5529–5539, 2014.

[13] J. Li, J. M. Bioucas-Dias, and A. Plaza. Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning. *IEEE Trans. Geosci. Remote Sens.*, 48(11):4085–4098, 2010.

[14] M.-A. Moen, A. P. Doulgeris, S. N. Anfinsen, A. H. Renner, N. Hughes, S. Gerland, and T. Eltoft. Comparison of feature based segmentation of full polarimetric SAR satellite sea ice images with manually drawn ice charts. *The Cryosphere*, 7(6):1693–1705, 2013.

[15] A. Qin and D. A. Clausi. Multivariate image segmentation using semantic region growing with adaptive edge penalty. *IEEE Trans. Image Process.*, 19(8):2157–2170, 2010.

[16] G. Schwarz et al. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.

[17] J. Shotton, J. Winn, C. Rother, and A. Criminisi. Textonboost: Joint appearance, shape and context modeling for multi-class object recognition and segmentation. In *ECCV*, pages 1–15. Springer, 2006.

[18] L.-K. Soh, C. Tsatsoulis, D. Gineris, and C. Bertoia. Arktos: An intelligent system for sar sea ice image classification. *IEEE Trans. Geosci. Remote Sens.*, 42(1):229–248, 2004.

[19] W. R. Tobler. A computer movie simulating urban growth in the Detroit region. *Economic geography*, pages 234–240, 1970.

[20] Q. Yu and D. A. Clausi. IRGS: image segmentation using edge penalties and region growing. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(12):2126–2139, 2008.

[21] Q. Yu, D. A. Clausi, et al. SAR sea-ice image analysis based on iterative region growing using semantics. *IEEE Trans. Geosci. Remote Sens.*, 45(12):3919, 2007.

[22] X. Zhu. Semi-supervised learning literature survey. Technical Report 2, University of Wisconsin-Madison, 2006.