# A New Look Into Image Classification: Bootstrap Approach

by

Shuhratchon Ochilov

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

Waterloo, Ontario, Canada, 2012

# Abstract

Scene classification is performed on countless remote sensing images in support of operational activities. Automating this process is preferable since manual pixel-level classification is not feasible for large scenes. However, developing such an algorithmic solution is a challenging task due to both scene complexities and sensor limitations. The objective is to develop efficient and accurate unsupervised methods for classification (i.e., assigning each pixel to an appropriate generic class) and for labeling (i.e., properly assigning true labels to each class). Unique from traditional approaches, the proposed bootstrap approach achieves classification and labeling without training data. Here, the full image is partitioned into subimages and the true classes found in each subimage are provided by the user. After these steps, the rest of the process is automatic. Each subimage is individually classified into regions and then using the joint information from all subimages and regions the optimal configuration of labels is found based on an objective function based on a Markov random field (MRF) model. The bootstrap approach has been successfully demonstrated with SAR sea-ice and lake ice images which represent challenging scenes used operationally for ship navigation, climate study, and ice fraction estimation. Accuracy assessment is based on evaluation conducted by third party experts. The bootstrap method is also demonstrated using synthetic and natural images. The impact of this technique is a repeatable and accurate methodology that generates classified maps faster than the standard methodology.

# Acknowledgements

I would like to express my appreciation to Prof. David Clausi for his supervision throughout this endeavor. I can not thank him enough for his persistent guidance and assistance that helped me improve my technical skills and academic knowledge.

I would like to express my gratitude to Prof. Melba Crawford of Purdue University for serving as my external examiner. I owe a debt of gratitude to my committee members Prof. Paul Fieguth, Prof. Claude Duguay, Prof. Zhou Wang for their invaluable feedback. I greatly appreciate the time they devoted and their sincere advices.

I greatly appreciate the support and co-operation of Prof. Claude Duguay. I have learned a lot from him through our discussions. I am also deeply grateful to Prof. Paul Fieguth for his inspirational course Statistical Image Processing.

I would like to thank Nic Svacina and Thomas Zagon for their help in conducting unbiased performance evaluation. It was pleasure for me to collaborate with them.

Thank you all to my dearest friends and brain-mates. It is their endless friendship, moral support and inspirational advice that I will never ever forget throughout my life.

I extend my deepest heartfelt thanks to my parents Zamira and Olim and to my wife Malika for their faithful support and confidence in me. I would always be indebted to them.

# Dedication

This work is dedicated to my dear mother Zamira, my dear father Olim, to my lovely
daughters Kamila and Rayhona, to my wife Malika and to rest of
important people in my life

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

A human can easily differentiate colors, objects, patterns from scenes, can understand and recognize sounds, characters, tastes, smells using past experience, can learn new knowledge by relating information from many sources. These acts exemplify classification, which a human can perform perfectly given sufficient resources, manageable quantity of samples and time. For decades, ongoing research has attempted to imitate the ability of the human to classify by mimicking human capabilities using computer algorithms. For instance, automated image classification, defined as the complete process of assigning qualitative labels for every pixel in the image [4] remains a key challenge in the image processing field [5, 6].

Availability of large image databases and the requirement for fast and accurate processing demand the use of automated image classification techniques [7]. Whether the image classification technique is unsupervised or supervised to some extent [8, 9] its aid is significant compared to manual classification of thousands of pixels. The results of image classification can be the basis for many environmental and socioeconomic models [10]. Classification of medical images, natural scenes and remotely sensed images are typical examples.

## 1.1 Challenges of Classifying Difficult Scenes

In different applications, various factors affect scene classification performance. Great effort has been made in developing novel techniques to improve classification accuracy [10]. However, classification challenges remain. Factors that impact the classification results include the complexity due to the nature of the image sensing, the size of the scene, the presence of many classes, the variability in class statistics. Generally, classification techniques can be grouped as supervised (Section 2.6.1) and unsupervised (Section 2.4) [10] methods, however in difficult scenes there is no known consistent method to yield high performance (Section 2.6.3) [7]. Especially, there is no full system defined that accounts for challenges in difficult scenes yet provides the complete classification.

Difficult scenes can be natural scenes containing complex objects with similar class statistics, can be medical images [8] (MRI [11], CT [12], ultrasound [13]) or can be remote sensed images [10] (multispectral [14], hyperspectral [15], sonar [16]).

Remotely sensed synthetic aperture radar (SAR) image is an example difficult scene. SAR provides additional features to optical sensors and is widely used in many applications. Remote sensing, particularly satellite SAR imaging, is used to study behavior and change of ice on the surface of sea and lakes [17, 18]. Target detection [19, 20], classification of urban [9,19–21] agricultural and forest images [9,22–25] are also common applications using SAR. Among broad application areas of SAR the challenges of sea and lake ice classification have been used as the example in the thesis.

## 1.2 Ice Classification

### 1.2.1 Why classify SAR ice images

**Sea-ice:**

Decrease in extent of sea-ice in polar regions and the trend of global warming provide clear evidence that the interest in sea-ice monitoring applications will increase. Subsequently, fast processing and interpretation of large volumes of wide expanse satellite data is required. Operationally, SAR images are now manually interpreted by assigning ice types and their concentrations to manually drawn large regions. An automatic procedure that assigns ice type labels to pixels is preferable to generate ice concentration maps.

The drawback of manual interpretation is that ice concentrations are often inaccurate. For example, a region may have 30% thick multi-year ice, but the user can not pinpoint the multi-year ice location. Such information would be useful for 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.

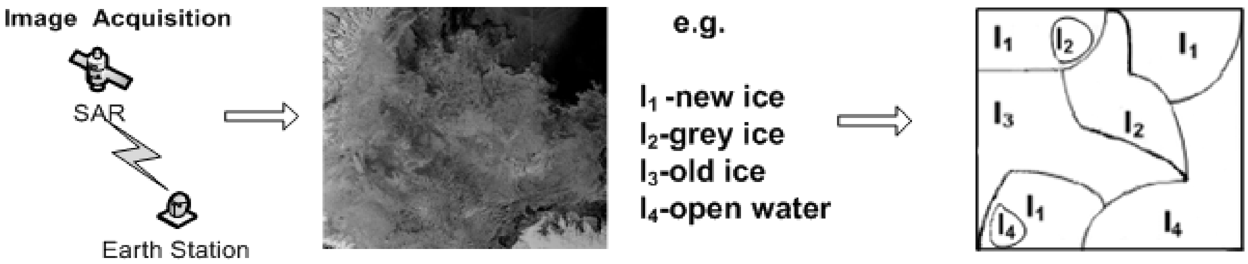
**Lake ice:**

Lake ice indicates the general trend of climate change when monitored consistently over a long period of time [26]. The evidence of climate change is generally related to the hydrology of northern regions, especially of big lakes such as Great Bear Lake (GBL) and Great Slave Lake (GSL). *In situ* measurements [27, 28] and analysis of remotely sensed images [29, 30] indicate that the duration of ice on the surface of the Canadian lakes is decreasing on average. Also, lake phenology parameters are the better descriptors of local climate characteristics when compared with air temperature [31]. Ice fraction, a lake phenology parameter, determines the expected precipitation and weather condition because of its significant role in shaping thermodynamic flow between the lake and the atmosphere [32]. Thus, if accurately estimated, ice fractions can be effectively used in weather and climate models.

### 1.2.2 Challenge of sea-ice classification

Independently, some of the component have been developed to partially address challenges of sea-ice classification but no system exists that provides the complete classification of difficult sea-ice scenes as depicted in Fig. 1.1. The contribution of the thesis is an end-to-end operational process which provides pixel-level classification of full SAR sea-ice imagery as shown in Fig. 1.2. In current processing flow, an ice analyst manually divides the image acquired by a SAR satellite(Fig. 1.1(a)) into large "polygon" regions (Fig. 1.2(a)) 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 identifying the spatial location of each ice type in the polygon region. To automatically interpret images based on egg code data, unsupervised classification can be performed on each polygon independently (Fig. 1.2(b)). Regions are automatically assigned a sea-ice label across the entire scene as shown in (Fig. 1.2(c)).

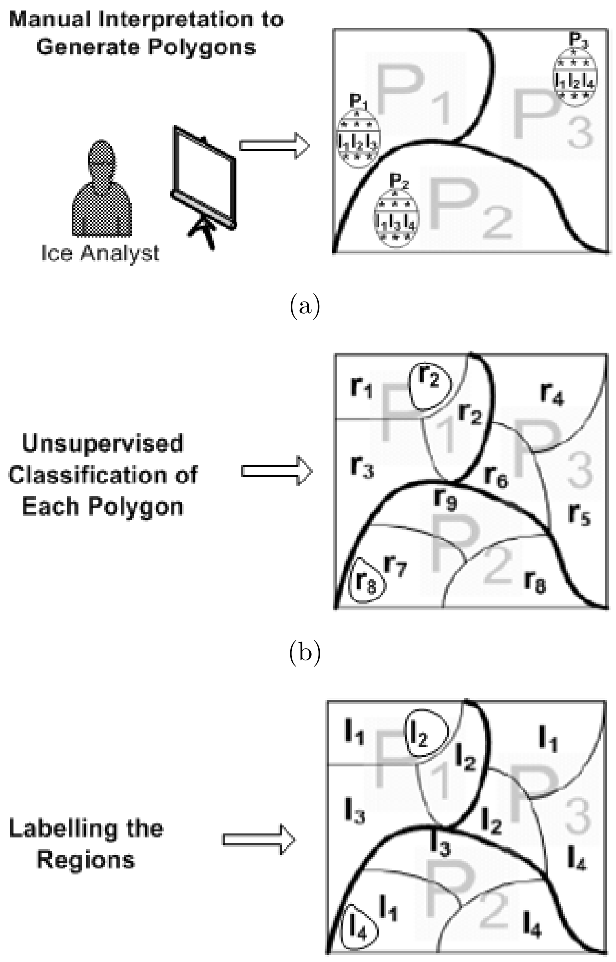


(a) (b)

Fig. 1.1: Classification problem of operational SAR sea-ice images. (a) SAR image acquired by satellite. (b) Example of classified SAR sea-ice image with qualitative labels assigned to every pixel.

### 1.2.3 Challenge of lake ice classification

At present, the Canadian Ice Service (CIS) provides end users with weekly ice fraction estimates of more then one hundred lakes across Canada and border of the United States. Limited time available for ice operators to process this data, ice fraction estimates are not always accurate. The experts at CIS visually investigate each SAR lake ice image and other sources (Section 2.2) to assess ice fractions. Furthermore, with the advent of new satellite missions the need to more frequently process data becomes evident which requires faster and automated methodologies. Thus, the same challenge described in previous Section 1.2.2(Fig. 1.1) is valid for lake ice classification. There is no classification system known that address the complete classification of difficult SAR lake ice images. Moreover,



(a)

(b)

(c)

Fig. 1.2: Processing flow of classification of full SAR sea-ice imagery. $r$, $P$, $l$ refer to sea-ice region, polygon and label. (a) Image manually divided into polygons with appropriate egg code data. (c) Image with every polygon automatically divided into $n_r$ regions using IRGS [2]. (d) Image with every region automatically labeled with sea-ice type. Thesis focuses on overall system for classification (a-c).

the components, such as unsupervised classification of full or partial lake ice imagery, are not widely studied. Previous studies applied to sea and river ice are discussed in Section 2.4. Methods [2,17,33] in SAR sea-ice interpretation field exist which have never been applied to SAR lake ice imagery.Thesisinvestigates the potential of classification system as described in Section 1.2.2(Fig. 1.2) that can be used for lake ice classification.

### 1.2.4 Challenge of validation

One of the main obstacles in developing algorithms for sea-ice interpretation is the lack of reference data. Fully validated field reference for the operational SAR sea-ice image is not available. For validation, one would have to perform field sampling of the ice, on site, across 500km by 500km region during the SAR satellite overpass. Such an exercise is logistically impossible and alternative methodologies need to be developed.

## 1.3 Thesis Contributions

### 1.3.1 Bootstrap approach for image classification

The main focus of this thesis is to present a single classification system which assigns qualitative labels to pixels in a rigorous, repeatable manner. An end-to-end framework, called a "bootstrap approach", has been developed to label the pixels which introduces the novel classification. By bootstrap approach the classification system is inferred which requires minimal input to start the process and the rest of the processing is fully automated. The following are the steps of the bootstrap approach (Fig. 1.3):

1. The user or automated technique separates the scene into non-overlapping subimages each with a constrained set of labels. Recently, hybrid approaches have demonstrated that for difficult scenes, some degree of user interaction or a constraint during the classification process can produce improvements [34–45] in unsupervised classification component of classification where the pixels are grouped into the regions. In bootstrap approach, such user input is also used to infer the qualitative labels for all pixels. Tessellation of the image into rectangles is an example of fast and simple generation of subimages and the "polygon", widely used terminology in the remote sensing literature, is an example of the subimage generated by user. For simplicity, the terminology polygon will be used for all subimages.

2. The next step is to perform unsupervised classification on each polygon using a most suitable technique for the application. Automatically classifying each polygon independently is an easier solution compared to unsupervised classification of a full scene. The unsupervised classification is more accurate when dealing with fewer classes having more stable local statistics.

3. Dividing the image into polygons increases number of regions produced by unsupervised classification, therefore, manual labeling of regions becomes time consuming. Automatic labeling technique has been developed which uniquely models the spatial relationship of regions between the polygons in the form of new neighborhood system embedded in a Markov random field (MRF) [46] framework.

4. The statistical and spatial relationship models are then combined in a Bayesian framework where the region labeling is formulated as a constrained optimization problem.

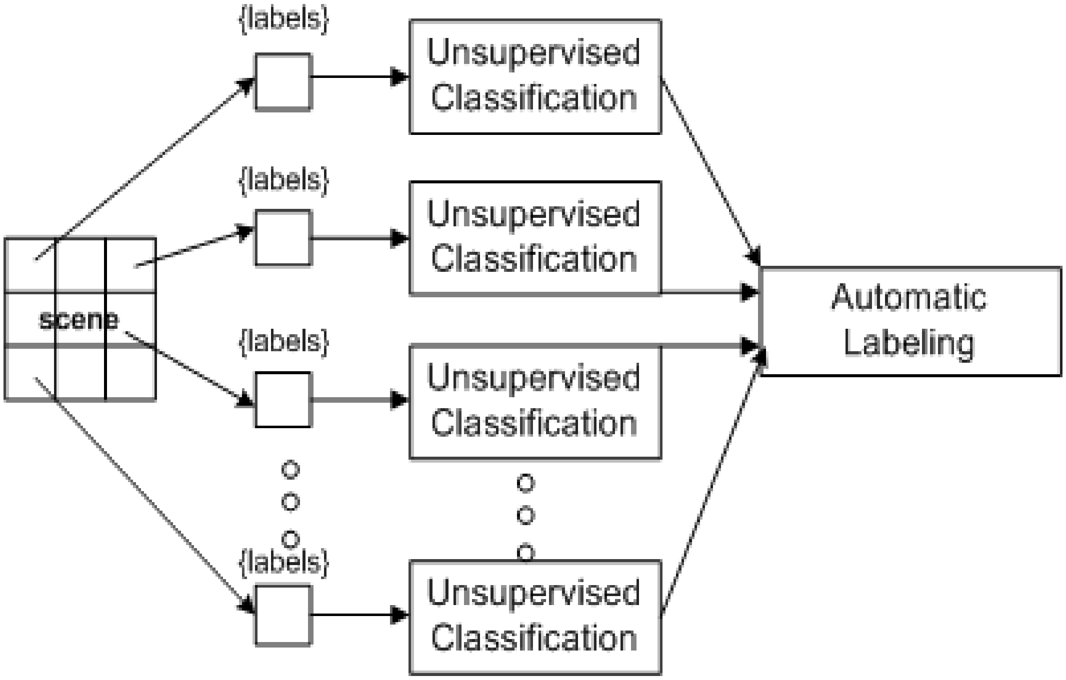


Fig. 1.3: Bootstrap approach for classification.

### 1.3.2 Bootstrap approach applications

**Generic Images:**

The automatic image labeling concept of bootstrap approach has been demonstrated in Section 4 with the generic images of nature and the synthetic scene. The images have been selected with different number of classes, class characteristics. The results indicate the success of the bootstrap approach where the final classification outputs contain the qualitative labels of the pixels (bird, grass, sky, mountain etc.). Robustness of the bootstrap approach have also been demonstrated in the presence of illumination distortion and additive noise.

**Sea-ice:**

Other research has considered the classical classification approach using training test samples [47–51]. The classification of SAR sea-ice images in Section 5 is performed in a novel way using the bootstrap approach. The SAR sea-ice image is divided into polygons as described in Section 1.2.2. Consequently, the polygons are automatically classified into disjoint regions using IRGS [2] where each polygon has its own set of class labels. Then, a global approach using all polygons is implemented to automatically label the regions as described in Section 1.3.1 which leads to an automatic algorithm for SAR sea-ice imagery which has not been accomplished before.

**Lake ice:**

There are techniques [2, 17, 33] used in the SAR sea-ice interpretation field that have not yet been applied to SAR lake ice. The characteristics and evolution of lake ice differs that from sea-ice (see Section 6.3). Hence, thesis initially investigates the potential of automatic algorithms that can be used for lake ice monitoring. For this purpose iterative region growing using semantics (IRGS) [2] and other conventional unsupervised classification algorithms are applied to the SAR images of GBL/GSL and their performance is compared. Ice fraction estimates are obtained over GBL/GSL to compare with CIS figures using the traditional classification approach. Introducing bootstrap approach for SAR lake ice classification (Section 6.5.3) has significantly increased the overall classification accuracy of difficult SAR lake ice scenes and has provided the automatic way to classify pixels as ice/no-ice.

### 1.3.3 Performance evaluation framework for validation

The most effective validation is performed by trained ice experts. As such, a user-interactive performance evaluation framework has been developed (Section 5.1.2) whereby a third party can validate both the unsupervised classification and labeling algorithmic outcomes. This is used to both evaluate the herein developed techniques and generate validated reference images.

# Chapter 2

# Background

Among broad applications of the bootstrap approach background is presented with examples from SAR sea-ice and lake ice. This chapter provides a brief background of physical aspects of SAR imagery and satellites used for ice monitoring. Application related specifics on SAR image understanding and data processing are provided in corresponding chapters. The methodologies for image classification are presented with the examples from SAR ice image classification. The system level classification approaches and the proposed bootstrap approach are described at the end.

## 2.1 SAR Fundamentals

Radar is an acronym for radio detection and ranging and in this methodology the microwave transmitter of a platform sends pulses to the sensed object which are scattered back and reach the receiver of the platform with the time delay $\Delta t$ as depicted in basic geometry in Fig. 2.1. In microwave systems, the essential material property is the relative dielectric constant $\varepsilon_r$ which is the measure of electrical energy conductivity. The time delay differences provide the resolution between different scatterers. The resolution in range and azimuth directions are independent. The resolution in the azimuth direction is the footprint of the antenna on the ground, and to obtain fine resolution, the large antenna is required.

To solve this problem the movement of sensing platform with a single antenna is sim-

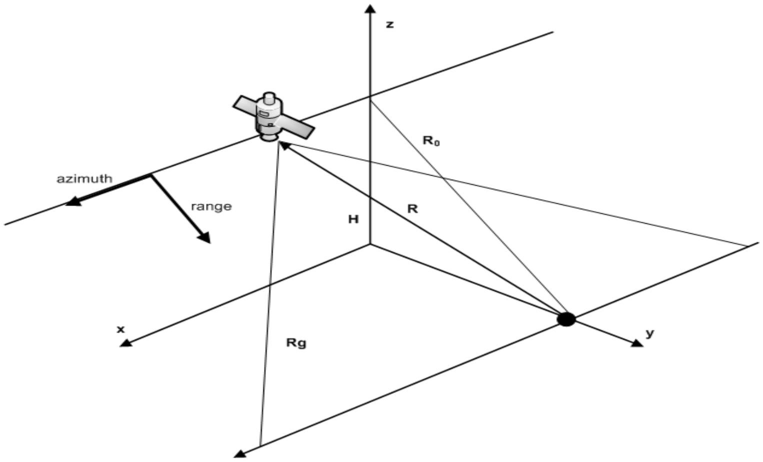


Fig. 2.1: Spaceborne SAR basic imaging geometry [52].

ulated as the array of sequential antennas along track [52]. The basic idea is that the scattered signal is contained in more then one pulse and exhibits as the phase history over the sensing period. Simulated sequential antennas are then combined in a single synthetic aperture by coherent combination of pulses. Thus, in SAR the physical length of antenna is substantially reduced and is only two times of the resolution required [52].

### 2.1.1 Satellites for ice monitoring

The use of optical sensors in northern regions is limited due to the polar darkness and extensive clouds during the freeze up period. This can be seen from Terra images of MODIS [53] covering the area of GBL/GSL; therefore, active and passive microwave remote sensing has been mostly used for ice monitoring [54–59].

Sensors used for ice studies are SeaWinds/QuikSCAT [59], special advanced microwave scanning radiometer - earth observing system (AMSR-E) [58], passive microwave and SAR. The QuikSCAT mission has been terminated at the end of 2009 and no replacement is expected to be launched in next five years [60]. AMSR-E is measuring the emitting radiation from the footprint of 76 by 44 km at 6.9 GHz to 6 by 4 km at 89.0 GHz [61]. Due to the low level of microwave radiation the passive microwave sensors have to integrate the incoming signals over large areas. Such a coarse resolution definitely does not provide sufficient resolution for strategic sea-ice mapping but useful for climate monitoring.

SAR is an active microwave sensor, provides finer resolution compared to passive microwave sensors and has ice penetration ability. These features make SAR crucial for sensing polar regions. Historically, aerial SAR has been used for strategic ice monitoring. 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 and Environmental Satellite (ENVISAT) the shift has been towards the use of spaceborne SAR ice imagery because of large coverage area and reduced costs.

RADARSAT-1,2 are commercial SAR satellites launched in 1995 and 2007 with the primary goal of monitoring and managing 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) [62] processing approximately 4000 SAR sea-ice images annually [63]. CIS is an organization responsible for providing ice conditions of Canadian shores, seas and lakes, in a timely manner, to support safe maritime operations [62]. Both satellites operate in 5.3GHz C-band suitable for ice type discrimination [64, 65]. RADARSAT acquisiton in ScanSAR Wideband mode is the combination of 50m fine spatial resolution with the coverage area of 500km. Although this is the preferred combination it introduces new level of complexity which is the processing of 10 K by 10 K images in a reasonable time. Manual pixel-level processing is not feasible and the requirement for automatic interpretation algorithms is obvious.

## 2.2 Data and Processing

### 2.2.1 RADARSAT-1 data

RADARSAT-1 images acquired in ScanSAR mode are the main source of SAR sea-ice and lake ice imagery used in this research. The archived data are sparse in some years due to availability of satellite acquisitions. The RADARSAT-1 ScanSAR Wide mode datasets provided by both the Canadian Ice Service and Canadian Space Agency (CSA) consist of five files: 1) volume directory file, 2) SAR leader file, 3) SAR data file, 4) SAR trailer file and 5) null volume directory file which have SAR sensing parameters, raw image values and ground control points (GCP). CSA provided images are block averaged to reduce 10

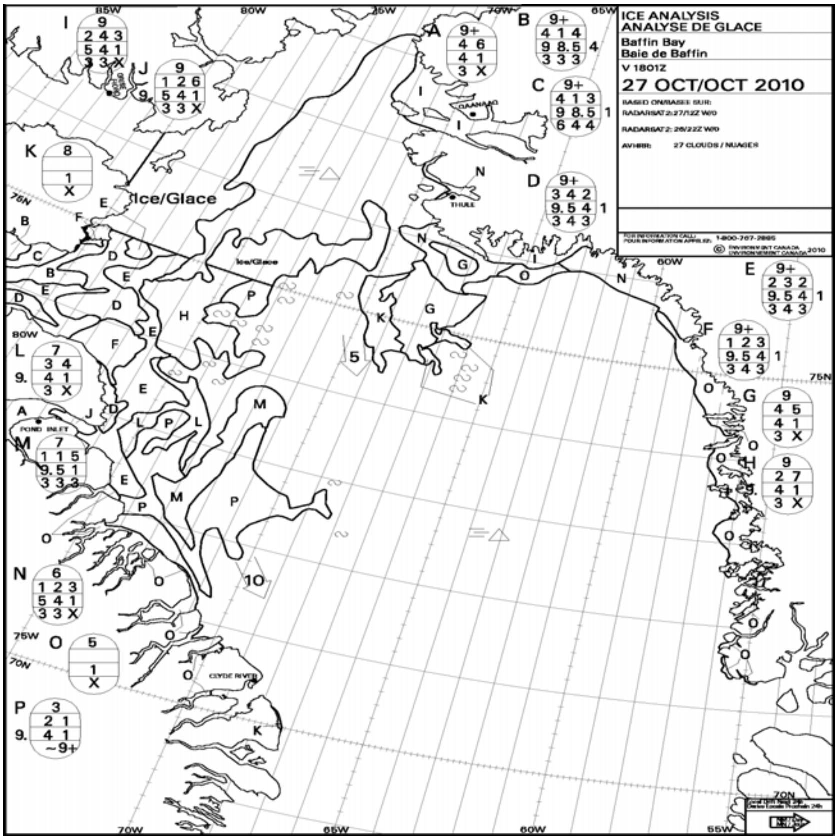


Fig. 2.2: The example of ice map provided by CIS (reduced to fit). Alphabet letters (A,B,C,..) associate egg code [1] with polygon regions.

Kx10 K images to 5 Kx5 K to be consistent with CIS data. CIS block averages SAR images for archival purposes with an assumption that the interpretation of 2x2 block average data is similar to the interpretation of full resolution data.

## 2.2.2 CIS ice maps and egg codes

Using the RADARSAT-1,2 SAR sea-ice images and ancillary information from ships, aircraft and meteorological sensors, a CIS analyst manually produces ice maps with egg codes and polygons similar to the map shown in Fig. 2.2. The standard sea-ice symbol representation along with descriptions are given in Fig. 2.3 and Table 2.1 per WMO standards [1].

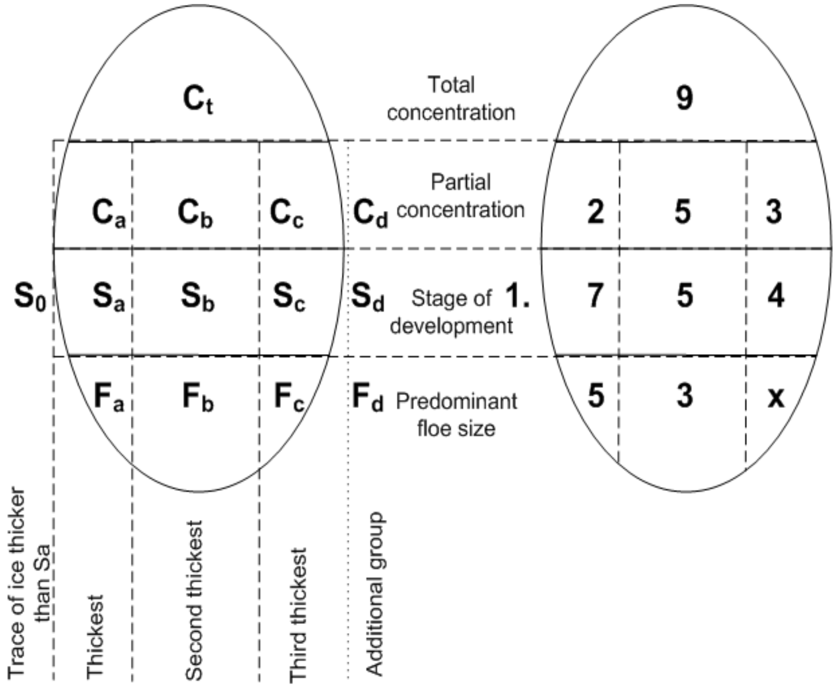


Fig. 2.3: The representation of sea-ice symbols in the sample egg code used by CIS per WMO standards [1].

### 2.2.3 CIS ice fraction data

CIS data contains ice fraction estimates of GBL/GSL and other lakes from year 1995 to present based on available RADARSAT images. The fractions, recorded in tenths, are manually assigned by an ice analyst who visually assesses RADARSAT scenes along with other sources. The data are reported weekly and dated on Fridays although the image under investigation might have been acquired earlier during the week.

### 2.2.4 Shape data

To accurately estimate lake ice fractions and to exclude the pixels with land and islands the lake boundary must be correctly located. The agencies providing the data do not provide the spatial interpretation of lake ice imagery. In this research the ArcGIS shape files have been retrieved from National Hydro Network (NHN) database [66]. In this manner, the features corresponding to islands have been excluded from further processing. Afterwards,

Table 2.1: The description of stage of sea-ice development $S_0, S_a, S_b, S_c, S_d$ per WMO standards [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 |

the conversion has been performed from ArcGIS proprietary .shp format into the MAGIC readable format and the lake boundaries have been automatically located.

### 2.2.5 MAGIC software system

MAGIC [3] is a software system to interpret SAR sea-ice images. MAGIC reads in the BIL and GCP files, and provides the visualization of RADARSAT images and subsequent analysis within an an easy-to-use GUI [3]. MAGIC [3] uses IRGS [2] as its core unsupervised classification algorithm. The key contribution of the thesis is the performance evaluation framework (Section 5.1) which has been implemented in the MAGIC.

Furthermore, MAGIC is able to read ASCII files consisting of polygon vertices in longitudes and latitudes which it projects to image coordinates. Projection is automatically performed using the Lambert Conformal Conic (LCC) projection [67] and polynomial fitting [68]. To date, MAGIC has focused on sea-ice, but its interface and algorithm capabilities are suitable for other images as well. In this thesis, we have provided extension of MAGIC to lake ice processing. The effectiveness is that to reach the pixel level accuracy MAGIC automatically masks out the land and internal islands by defining accurate boundaries of lakes from their geographical location data (Section 2.2.4). Thus, the automated unsupervised classification by IRGS only includes the surface of the lake.

## 2.3 Feature Extraction

Important step in a pattern recognition algorithm is the extraction of meaningful features. The features extracted from sea-ice images can be grouped as tone, texture or shape. Features maybe are extracted at different stages of classification process. The following sections explain the tone, texture and shape feature extraction methodologies used in the literature for ice classification.

### 2.3.1 Backscatter

In SAR remote sensing the radar cross section (RCS) $\sigma$ is measured using the relationship between the transmitted power and the power of the backscatter from the sensed object received at the antenna on the platform [69]. Subsequently, the “sigma nought” $\sigma^0$ is computed by normalizing RCS by the area:

$$\sigma^0 = \frac{\sigma}{A} \tag{2.1}$$

and, generally, expressed in decibels in the range of [-5dB to -40dB] [70]. For an 8 bit image the range of $\sigma^0$ is quantized into 256 levels.

### 2.3.2 Shape features

In addition to backscatter, the features can be extracted from SAR sea-ice images related to shape of objects in the scene. Floes and leads are two main objects that have defined structure and their shape characteristics convey the information about the ice-type of the region. There is no indication in the literature of using shape features for unsupervised classification of SAR sea-ice images; however, shape feature have been utilized for sea-ice labeling by Soh et al [71] and Yu and Clausi [33], where shape features are extracted from the regions provided by unsupervised classification. Those techniques may work for well-defined shapes with clearly identified boundaries but in SAR sea-ice images, the shape boundaries are generally obscure and small floes/leads that are close together can be combined by shape detectors into a single irregular shape which may deceive the labeling technique.

### 2.3.3 Texture features

Texture itself has many definitions, for instance, Petrou and Sevilla have described texture as a local variation and dependence of data at lower scales [72]. There many statistical, model based, geometrical and domain transform based techniques exist for extracting features from textures such as co-occurrence matrices, fractals, morphology, power spectrum, autocorrelation, wavelet transform, Gabor functions [72]. The two texture descriptors that have been extensively studied and used in SAR sea-ice classification applications are probabilistic grey-level co-occurrence matrix GLCM [47, 48, 51, 73–75] and signal processing based Gabor [49, 76, 77]. For many years the great emphasis has been given to texture features as a main descriptor of sea-ice classes, however in difficult SAR scenes classification with texture features face the same challenges as described in Section. 2.6.1.

## 2.4 Unsupervised classification

The task of unsupervised classification methods is to automatically separate the image into disjoint regions having sites with some common characteristics. The site could be pixel or watershed region [78] and how their commonality is defined and modeled is what categorizes the different unsupervised classification methods.

Due to sensor and environmental limitations (Section 2.4.1) SAR image unsupervised classification is considered as nontrivial.

The mathematical formulation of the unsupervised classification problem is given in Section 2.4.2. In subsequent sections, the four techniques are presented, which are the representatives of the majority approaches of unsupervised classification methods. Those techniques are pixel-based K-means, pixel-based Gaussian mixture model (GMM), region-based constant multi-level logistics (C-MLL) and region-based IRGS [2]. Pixel-based or region-based refers to unsupervised classification which treat sites as pixels or watershed regions, respectively. The initial oversegmentation into watershed [78] regions reduces the processing time considerably, therefore highly advisable for MRF techniques such as C-MLL and IRGS.

### 2.4.1 Challenges of SAR image unsupervised classification

#### Sensor effects

Speckle noise: One of the phenomena in SAR images is the effect of speckle, generally, modeled as multiplicative noise [9, 23, 79, 80]. Due to speckle noise, there is a high intra-class variability causing SAR images to appear grainy with details such as edges and objects degraded [80]. One approach to solve this problem is the use of speckle denoising techniques [81–83] which unavoidably smooth image details [80]. Another approach is to use a classification technique which is robust to speckle noise. The conventional classification techniques which treat sites independently are in disadvantage for such requirement.

Incidence angle: The incidence angle, is one of the contributors to nonstationarity in SAR imagery [84, 85]. The incidence angle effect is more evident in large scenes where characteristics of classes changes within the scene significantly. This results in high inter-class variability of the backscatter.

#### Environmental effects

Surface roughness and environmental conditions are key contributors to nonstationarity in SAR imagery. For example, wind may result in roughed water which exhibits different characteristics from calm water. This also results in high inter and intra-class variability of the backscatter.

### 2.4.2 Formulation of unsupervised classification

The problem of unsupervised classification of any image can be formulated as follows. Let $S$ denote a discrete 2D rectangular image space of size $M \times N$. $Y = \{Y_s | s \in S\}$ represents the 2D random variable defined on $S$. Suppose there are $n$ different regions in $Y$. Let $X = \{X_s | s \in S\}$ be another univariate 2D random variable defined on $S$ which is related to the original image $Y$ where each discrete valued random variable $X_s$, having a value in $\{1, \ldots n\}$, represents the regions to which the site $s$ belongs. Suppose the realization of $Y$ and $X$ are $y = \{y_s | s \in S\}$ and $x = \{x_s | s \in S\}$ respectively, then the image unsupervised classification problem can be formulated as an estimation of $x$ from $y$:

$$R : \{y_s | s \in S\} \longrightarrow \{x_s | s \in S\} \tag{2.2}$$

based on some objective function. Subsequently, the unsupervised classification result $X = x$ consist of $n$ disjoint regions $\{\Omega_1, \Omega_2, \ldots, \Omega_n\}$ such that

$$\begin{aligned} &1. \quad \Omega_i = \{s |\, X_s = i, s \in S\} \\ &2. \quad \forall\, i \neq j, \Omega_i \cap \Omega_j = \emptyset \\ &3. \quad S = \bigcup_{i=1}^{n} \Omega_i \end{aligned} \tag{2.3}$$

where $1 \leq i \leq n$ and $1 \leq j \leq n$ are indices of regions [74]. All four methods in next sections are trying to solve the Eq. 2.2 but using different objective functions.

### 2.4.3 Non-spatial clustering methods

#### K-means pixel-based

K-means is a non-spatial unsupervised classification method. By non-spatial, the methods are defined that do not account for the spatial interactions between the sites which are first separated into clusters and then placed together to achieve image unsupervised classification formulated in Section 2.4.2. K-means automatically classifies the image by iteratively optimizing the following objective function [86]:

$$\arg \min_{\{x_s | s \in S\}} \sum_{i=1}^{n} \sum_{s \in \Omega_i} (y_s - \mu_{\Omega_i})^2 \tag{2.4}$$

For the realizations $\{x_s|s \in S\}$, Eq. 2.4 basically tries to minimize the total least square error (LSE) between the pixel values $y_s$ and the mean of regions $\mu_{\Omega_i}$. As seen from Eq. 2.4 K-means does not utilize spatial proximity of sites which will deteriorate the performance in the presence of noise.

#### GMM pixel-based

The GMM [86] and K-means are similar since both of the techniques are non-spatial, therefore the effect of noise is still eminent in GMM. However, the GMM tries to iteratively fit the pixels into Gaussian mixture model. Such a distribution model helps in cases of high overlap in feature space since the decision metric is based on both the distance to the mean as well as covariance of the regions. This could be better seen from objective function given below:

$$\arg \min_{\{x_s|s \in S\}} \sum_{i=1}^{n} \sum_{s \in \Omega_i} \left\{ \frac{1}{2} \ln(2\pi)^K |\Sigma_{\Omega_i}| + \frac{1}{2}(y_s - \mu_{\Omega_i})^{tp} \Sigma_{\Omega_i}^{-1} (y_s - \mu_{\Omega_i}) \right\} \tag{2.5}$$

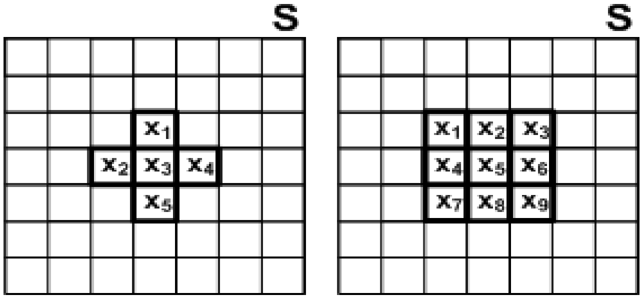
where $tp$ is the transpose operation, $\mu_{\Omega_i}$ $\Sigma_{\Omega_i}$ are the mean and covariance of region, $K$ is the dimension of feature vectors. In the thesis feature vector has the dimension of $K = 1$ and corresponds to the backscatter intensity (tone) of a pixel. Equation 2.5 is simply the expression of Gaussian distribution summed over all regions.

### 2.4.4 MRF-based unsupervised classification

#### Markov random field

A random field can be defined on $S$ with $X = \{X_s|s \in S\}$ being a set of variables which are the outcome of the random process, as an example the outcome can be the number of heads that occur from an unbiased coin tossing experiment [72]. However, in Markov random fields the coin is biased and the probability of some value occurring for a particular member of an image is dependent on its neighbors. That is to say, the neighbors of any image member $x_s$ do have the information about the rest of the image and knowing their values determines the conditional probability.

$$p(x_3|x_s, s \in S, s \neq 3) = p(x_3|x_1, x_2, x_4, x_5) \tag{2.6}$$



(a) (b)

Fig. 2.4: (a) MRF with first-order neighborhood. (b) MRF with second-order neighborhood.

This property is also known as a Markovianity principle and simple illustration is shown in Fig. 2.4(a) and expressed with Eq. 2.6.

### Region-based processing

To accommodate challenges of unsupervised classification of SAR images, the region-based techniques [2,9,23,24,80,87] have been presented in the literature which deal with a group of pixels (i.e., region) as a single entity (i.e., site). For instance, the group of pixels can be watershed region [2,87] or an element of polygonal tesselated image [23,24]. Using oversegmented regions as initial point reduces the processing time considerably, therefore highly advisable for the techniques which complexity is a function of sites.

### C-MLL region-based

C-MLL is a classical MRF based unsupervised classification model [88,89] approach which iteratively tries to minimize following objective function

$$\arg \min_{\{x_s|s\in S\}} \sum_{i=1}^{n} \sum_{s\in\Omega_i} \left\{ \frac{1}{2}\ln(2\pi)^K|\Sigma_{\Omega_i}| + \frac{1}{2}(y_s-\mu_{\Omega_i})^{tp}\Sigma_{\Omega_i}^{-1}(y_s-\mu_{\Omega_i}) + \beta \sum_{t\in N_s}\delta(x_s,x_t) \right\} \tag{2.7}$$

where $\beta$ is an empirically derived parameter, $N_s$ are the neighbors of site $s$, $\delta()$ is the Kronecker delta function. The Eq. 2.7 contains the GMM expression added to spatial

MLL model. The spatial model simply states that the neighboring regions are likely to be the members of the same region. Such a formulation reduces the effect of noise by forcing a formation of smoother regions. However, the static parameter $\beta$ in C-MLL model does not efficiently address the nonstationary in SAR images, therefore, may not be suitable in complex scenarios. One similar implementation of Eq.2.7 is varying MLL (V-MLL) differ in a way the parameters are specified. In V-MLL [90] the parameter $\beta$ is set to one and different parameter $\alpha$ is set to control the feature model by monotonically decreasing during the iterative process.

**IRGS region-based**

IRGS [2] uses the same objective function as in Eq. 2.7 but with a different spatial term which is based on edge penalty. Such spatial term is demonstrated in the objective function of IRGS:

$$\arg\min_{\{x_s|s\in S\}} \sum_{i=1}^{n}\sum_{s\in\Omega_i} \left\{ \frac{1}{2}\ln(2\pi)^K|\Sigma_{\Omega_i}| + \frac{1}{2}(y_s-\mu_{\Omega_i})^{tp}\Sigma_{\Omega_i}^{-1}(y_s-\mu_{\Omega_i}) + \beta\sum_{t\in N_s}(1-\delta(x_s,x_t))\exp\{-(\frac{\nabla_{st}}{K(\theta)})^2\} \right\} \tag{2.8}$$

where $\beta$ is the parameter estimated as in [91], $\nabla$ is the gradient between two neighboring sites, and $K(\theta)$ is the monotonically decreasing function of iteration $\theta$ [2]. In this case, two sites are likely to be assigned to the same region depending on both being neighbors and the gradient between them. The less the gradient, the greater the edge penalty imposed if neighboring sites are assigned to different regions. The nonstationarity and the noise in SAR images are addressed with a varying parameter $K(\theta)$ inversely affecting the spatial term. Hence, the spatial term is gradually increasing in IRGS as a function of iterations. In such a scheme the effect of the spatial term contributes less at few initial iterations, whereas, at later stages more weighting is applied on strong edges and the weak ones are prone to dissolve producing more definitive regions.

### 2.4.5 Other methods

The methods described until now are the examples from SAR sea-ice and lake ice unsupervised classification. This section briefly describes other methods available in the literature. The MRF-based methods are dependant on local neighborhoods and do not directly incorporate global features available in the image [45]. Thus, to address this fact and to incorporate the active inputs from user the multilayer Bayesian network has been proposed which incorporates the superpixel regions, boundaries, nodes and actively incorporates user intervention. Such methods will better perform for unsupervised classification of smaller scenes containing global objects but may not be as efficient for unsupervised classification of SAR sea-ice scenes which challenges have been mentioned in Section 2.4.1. Similar to MRF-based approaches are the graph theoretic approaches [92, 93] which have gained popularity over the last decade. In such a method [92] the objective function consisting of quality term (same as feature term) measures the quality based on some clustering technique (e.g. techniques in Section 2.4.3) and the spatial coherency term (same as prior term) which for example penalizes if the realization has neighboring pixels belonging to different classes. Consequently the objective function is solved using the optimization method based on graph cuts [94–96] which has been proven to best work with stereo images.

The nonparametric approaches in feature space have gained traction because of adaptability to different scenes and of no need to tune the parameters. MRF-based methods, generally, use such parameters. Simple technique to achieve segmentation is histogram thresholding. One of the earlier methods [97] calculated the threshold by minimize within class variance and the more advanced automatic thresholding is presented using fuzzy measures [98]. The histogram thresholding methods are mainly relying on grey-level intensities of pixels. A recent hierarchical method [99] incorporates the intensity texture and takes the geometry of the regions as a prior in probabilistic model. The number of classes and the probability density function p.d.f. do not have to be known for nonparametric segmentation methods based on mean-shift [100]. This methods still employs the global feature space information which make them prone to fail in noisy images. To accommodate the noise the spatial context is introduced in clustering technique, fuzzy local information C-Means [101], which is free of parameter selection and uses the soft decisions for the clustering.

The fusion of segmentation result from several simple technique is achieved with MRF-based combining [102]. The goal of the fusion is to eliminate the deficiency of simple segmentation techniques and the complexity of sophisticated and computationally expensive segmentation models.

## 2.5 Image Labeling

Despite progress in unsupervised classification field, limited research has been performed in qualitative labeling. The image labeling is discussed in the context of classification systems overview (Section 2.6).

## 2.6 Classification systems overview

### 2.6.1 Supervised approach

The traditional way of classifying images is supervised (Fig. 2.5). The user knows visually the pixel location of each class and manually selects samples for each class to create training samples for statistical analysis. The classifier parameters obtained from training samples are used to distinguish and to qualitatively label every pixel in the scene [103]. The practical and simple supervised classification technique is maximum likelihood classifier [10]. The advantage of supervised classification is that the user is involved only at the initial step and the rest of the process is automatic. However, the extraction of representative training samples is time consuming, especially, if the classes are spatially mixed. Furthermore, classification based on user-selected and previously obtained training samples is very sensitive to number of samples, spatial autocorrelation, intra and inter class variance [10, 17, 104–107].

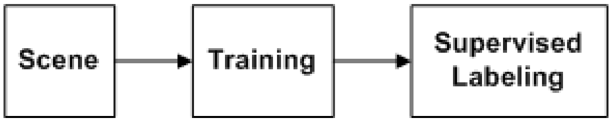


Fig. 2.5: Supervised approach for classification.

**Sea ice:**

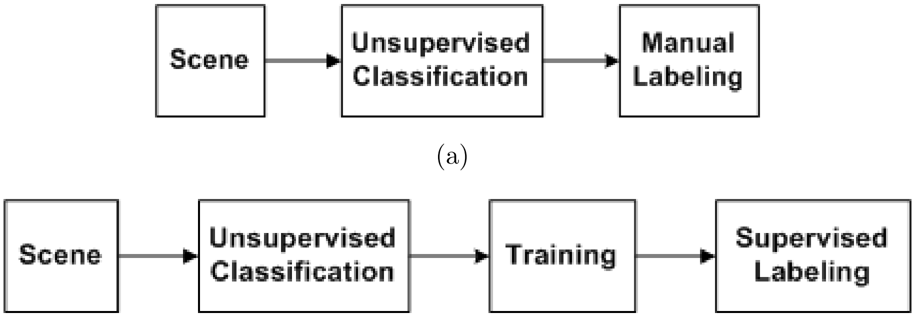
There is a history of supervised classification using windowed sample data. For example, numerous studies [47–51] follow the classical approach of carefully gathering training samples to extract class statistics. In this approach, the ice analyst cumbersomely discerns and selects pure ice samples from 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. Hence, for operational use, techniques that classify based on thresholds [50, 71, 108] 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. In an algorithmic sense, the classification can be achieved by sea-ice labeling using the features obtained from individual images following initial unsupervised classification.

**Lake ice:**

The Great Lakes winter experiment (GLAWEX'97) by Nghiem and Leshkevich [109] has been been conducted in the field of supervised SAR lake ice interpretation. The target of the experiment has been to obtain the signature library of the backscatter and corresponding physical properties of different ice types, on two dates over the Great Lakes. A C-band polarimetric scatterometer instrument on board of a ship was used to collect the data which has been used for classification of RADARSAT-1 and ERS-2 imagery of Lake Superior (1996 and 1997) [110]. As discussed in Section 6.3.2 the backscatter characteristics of ice varies at different dates and locations. Environmental conditions such as abrupt changes in air temperature, roughening of open water, snow cover and changing incidence angle affect backscatter characteristics. Operating with only training samples might deteriorate ice interpretation performance, thus the need of unsupervised classification algorithms is evident to extract features within the scene.

### 2.6.2 Unsupervised approach

In unsupervised classification the image is automatically classified into regions using features within the image. This approach is more data driven than supervised classification. In unsupervised classification, the image is classified into the regions which number can be specified by the user or inferred from data. If the unsupervised classification of a full scene is successful, which is not always the case for difficult scenes, the assignment of qualitative labels could be done in two ways as shown in Fig. 2.6. The user can interpret the image and manually assign a qualitative label to each region (Fig. 2.6(a)). Generally, the number of classes needs to be provided by user before the unsupervised classification, therefore in this way user will be involved at two stages of classification. The alternative way (Fig. 2.6(b)) is to provide the domain knowledge for supervised labeling of regions. The domain knowledge could be again the training data such as class statistics, thresholds which drawbacks have been already mentioned. The preferable way is to provide the domain knowledge which is data driven.



(a)

(b)

Fig. 2.6: Unsupervised approaches for classification.

**Sea ice:**

Published research [2, 3, 111–113] exists that classifies SAR sea-ice scenes using an unsupervised approach. One of the earlier works is presented by Samadani [112] where the unsupervised classification on airborne SAR images is achieved by modeling features using a Gamma mixture. Deng and Clausi [111] obtained pixel-based unsupervised classification using an MRF framework by extracting texture features from SAR sea-ice imagery, modeling them as a Gamma mixture and adding the spatial varying term which they initially proposed in [90]. Soh and Tsatsoulis [113] developed automated sea-ice unsupervised classification ASIS algorithm and ARKTOS [71] system for ERS 1,2 and RADARSAT imagery which automatically determines the number of classes and performs unsupervised classification. ARKTOS has been evaluated by CIS and deemed not appropriate for operational needs. Yu and Clausi [2] implemented a fast region based unsupervised classification algorithm which incorporates a gradually increasing edge penalty (GIEP) MRF model and region merging operation and referred to the whole system as iterative region growing using semantics (IRGS). Among different techniques, IRGS [2] has shown robust performance classifying operational SAR sea-ice imagery. IRGS has been successfully tested and validated by CIS personnel. For these reasons, IRGS has been used in the thesis for unsupervised classification of sea-ice.

**Lake ice:**

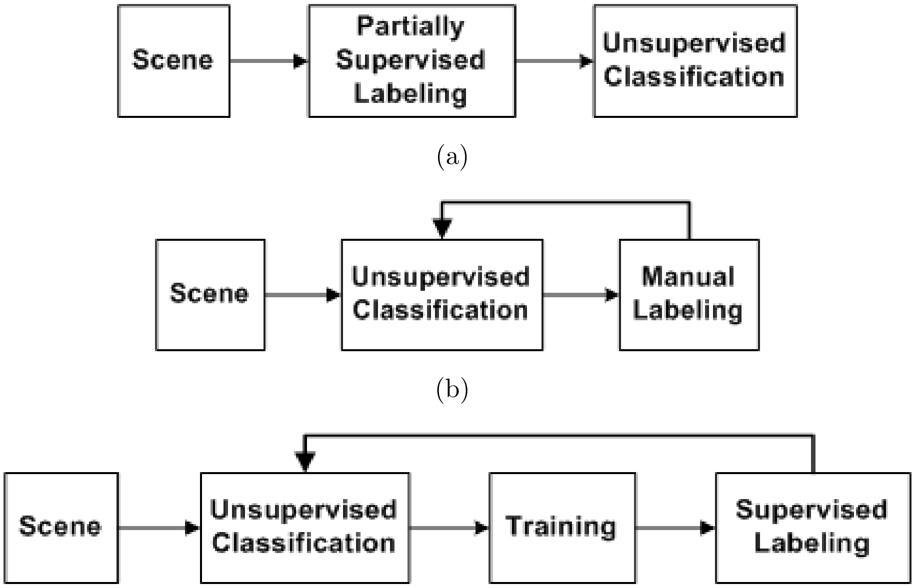
There are no validated unsupervised classification techniques known for SAR lake ice image interpretation. Automated techniques have been developed for river [114] and sea-ice applications but never have been applied to lake ice. The framework has been demonstrated to automatically monitor river ice using fuzzy K-means [114,115] unsupervised classification of texture features [116]. K-means and histogram thresholding [97], spatially unaware techniques, perform the unsupervised classification in the feature space and therefore will strongly be affected by noise.

### 2.6.3 Hybrid approach

The distinction of supervised and unsupervised techniques is becoming vague with the recent utilization of constrained and partially labeled data [41]. Since unsupervised classification is an ill-posed problem [117], for difficult scenes it may not provide meaningful results, therefore, any additional user provided information may help the unsupervised classification process to achieve a better solution [34, 35].

1. Partial labeling or grouping is a side information included in unsupervised classification process in the form of soft or hard constraints [34,40–42] (Fig. 2.7(a)). In partial labeling, sites are selected from the image and assigned a label or “must link” “must not link” pairwise constraints are imposed during the optimization. Additionally, the focus of unsupervised classification could be confined to the location of interest [40].
2. Relevance feedback provides user information to monitor and provide feedback on unsupervised classification outcomes [36, 118] (Fig. 2.7(b)). In such schemes, the unsupervised classification is used as an exploratory tool rather then single direction technique [36].

3. Yu and Clausi [33] classified SAR sea-ice images in an iterative manner based on an objective function adapted from Zhang and Modestino [119]. The authors recognized that the blind unsupervised classification with no guide of domain knowledge is prone to produce inaccurate unsupervised classification results which are further misclassified in the labeling stage. As opposed to conventional unsupervised approaches, the authors' classification approach is iterative where intermediate results produced in unsupervised classification stage of IRGS [2] are assessed by supervised labeling using the parameters and threshold derived from training (Fig. 2.7(c)).



(a)

(b)

(c)

Fig. 2.7: Hybrid approaches for classification. (a) The image is partially labeled as in supervised approach and passed as a constraint to unsupervised classification. (b) As in unsupervised approaches the user manually labels the unsupervised classification results but provides a relevance feedback to unsupervised classification to refine the results. (c) The training data is used to set the rules for automated supervised labeling which provides the feedback back to unsupervised classification to refine the results.

### 2.6.4 Summary and the proposed approach

Existing classification techniques [33, 71] either use supervised approaches or make use of unsupervised classification as their initial step. Labeling images especially SAR images, based on training statistics is not expected to be robust in an operational environment. The preferable way is to provide the domain knowledge which is data driven. If the unsupervised classification of a full scene is successful, the image can be labeled manually. This is not always the case for difficult scenes and to help unsupervised classification the user can separate the image into polygons, which will result in many regions (Fig. 1.3). The manual labeling of many regions can be time consuming. An initial attempt to avoid the training step [73] used a logical assignment of ice labels by cognitive reasoning and class statistics, but this outcome is dependent on the ordering of the polygons. However, labeling is preferred to be optimal over the whole scene, and can be performed automatically by the best fit modeling of information from all polygons as shown in Fig. 1.3.

# Chapter 3

# Methodology for Bootstrap Approach

This chapter builds the mathematical model for bootstrap approach (Section 3.1). A fully automated Markov random field model for bootstrap approach is described in Section 3.2 that is used to combine the initial labeling constraints and regions from unsupervised classification to produce fully classified full scene. The implementation scheme is described in Section 3.3 and the algorithmic flow presented in Section 3.4 following the strategy for applying the methodology in Section 3.5.

## 3.1 Mathematical Model

Let $S$ denote a discrete 2D rectangular image space of $M \times N$ number of pixels ($s$). In bootstrap approach the image space is partitioned into polygons. The polygon $P = \{P_1, P_2 \ldots P_{n_p}\}$ is a higher structure defined on image space $S$ (Fig. 3.1(a)). Based on application the most suitable automatic unsupervised classification is performed separately on each polygon where each polygon $P_q$ consists of one or more regions from the set $\{r_i, \ldots\} \subseteq r$ (Fig. 3.1(a)). There may be multiple segments assigned to the same $r_i$. Consider representing the image space $S$ as $n_r$ closed *regions* $r = \{r_1, r_2, \ldots, r_{n_r}\}$ having boundaries ($\partial$). In difficult scenes (Section. 1.1) the feature statistics may not solely represent the class characteristics, therefore, the interaction between neighboring polygons is important. The drawback of classification, which considers polygons independently with no spatial context, has been described in Section 2.6.4. Certainly, this drawback has motivated proposition of MRF model to combine the information from all polygons, regions from unsupervised classification and their spatial interactions to produce consistently classified full scene. The following expressions and definitions are necessary to build the model.

Assume the unsupervised classification result $X = \{X_r | r \in S\}$ represents the random field defined on $S$, where $X_r$ is the random variable representing all pixels $s \in r_i$. Suppose there are $n_l$ different classes in $X$. Let $Z = \{Z_r | r \in S\}$ be another random field defined on $S$ where each discrete valued random variable $Z_r$, has value in $\{1, \ldots n_l\}$, represents a label to which all pixels $s \in r_i$ in that region belong. Hence, the classification problem can be defined next.

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

$$L : \left\{ \; \{x_r | r \in S\} \;\; \longrightarrow \{z_r | r \in S\} \right. \tag{3.1}$$

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

The problem definition can be best visualized in Fig. 3.1. The task is to obtain fully classified scene in Fig. 3.1(b) using the unsupervised classification result in Fig. 3.1(a) along with the initial constraints as given in the same figure notes. The problem formulation is unique and if solved provides the methodology for bootstrap approach where the end-to-end classification is achieved by only providing the polygons with labeling constraints.

***Definition 2:*** The *boundary* is the set of all neighbor pixels between regions. The neighborhood of pixels is determined based on first-order neighborhood system [88](Section. 2.4.4) defined on image space S. If at least one such pixel exists the regions are said to have a boundary between them as depicted in Fig. 3.2(a).

Boundaries help define the neighborhood of regions.The determination of neighboring pixels is not limited to first-order and can be extended to higher-order neighborhoods. The novel neighborhood is defined next which encompasses the interaction between the polygons.

***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. $N_{r_i}$ 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 $r_j \in N_{r_i} \Leftrightarrow r_i \in N_{r_j}$. The *Neighborhood system* $N_r$ is a set of all neighborhoods.

In Fig. 3.1(a) $r_2$ and $r_4$ are neighbors but $r_1$ and $r_2$ are not as depicted in graphical representation in Fig. 3.2(b). Basically, the inter-polygon spatial interactions are focal, whereas, the intra-polygon spatial interactions are neglected since two regions in same polygon can not be assigned the same label during final classification (*Definition 1*).

***Definition 4:*** *Clique* $\mathsf{c}$ is a single region or a subset of regions for which every pair of regions are neighbors and the set of all cliques is $\mathsf{C} = [\mathsf{c} \mid \mathsf{c} \subset N_r]$ [88].

Example of first-order cliques is shown in Fig. 3.2(c).

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

$$p(z_r|S, r, s \notin S) = p(z_r|N_r) \tag{3.2}$$

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 conditionally decorrelated and independent ones [120].

***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 [88, pp. 28]

$$p(z_r) = \frac{1}{D} \exp\left\{ -\frac{1}{T} \sum_{\mathsf{c} \in \mathsf{C}} V_{\mathsf{c}}(x_r) \right\} \tag{3.3}$$

where $D$ is the partition function that normalizes the distribution, $T$ is the temperature,

$V_{\mathsf{c}}(x_r)$ is the clique potential and the negative of exponential $\sum_{\mathsf{c} \in \mathsf{C}} V_{\mathsf{c}}(x_r)$ is called clique energy $E$.

The theorem is the framework for computing the probability of an MRF [72].

## 3.2 Automatic Image Labeling

The clique energy model  is mainly based on following ideas:

- The feature model has to be prioritized at initial iterations to correctly estimate the parameters of classes. Thereby, the weight of data model can be a monotonically decreasing function of iteration.
- Interaction between regions in different polygons modeled as pairwise clique energies can have the form of edge penalty to accommodate two polygon generations. 1)Polygon is generated to separate regions with different classes. 2)Polygon is generated to divide large regions with the same class into smaller ones.
- Due to storage and processing limitations of large scenes the single and pairwise clique energy have to be defined in computationally efficient way otherwise the labeling model has less usability in real world applications. Thus, the pixel-based processing needs to be minimized.

Labeling is performed by maximum a posteriori MAP estimation. The Bayesian theory:

$$p(Z = z|X = x) = \frac{p(X = x|Z = z)p(Z = z)}{\sum p(X = x|Z = z)p(Z = z)} \tag{3.4}$$

states that *a posteriori* probability $p(Z = z|X = x)$ is equal to the product of likelihood $p(X = x|Z = z)$ and *a priori* probability $p(Z = z)$, divided by normalizing constant [121]. The optimal estimator (MAP) which maximizes the *a posteriori* probability can be given as:

$$\begin{aligned} &\arg \max_{\{z_r | r \in S\}} p(Z = z|X = x) \propto \\ &\arg \max_{\{z_r | r \in S\}} p(X = x|Z = z)p(Z = z) \end{aligned} \tag{3.5}$$

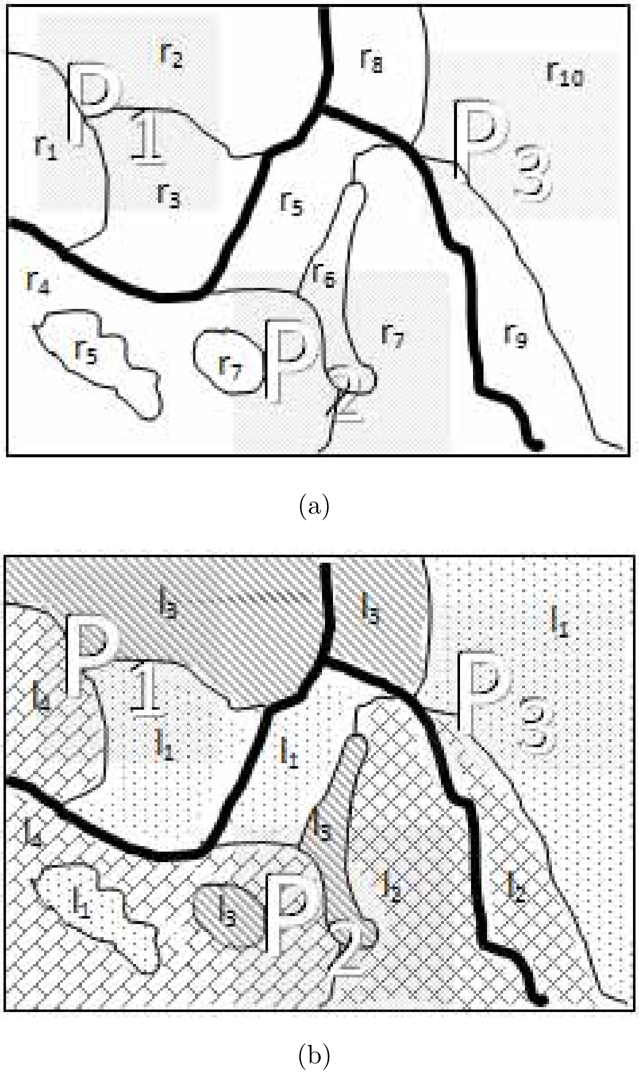
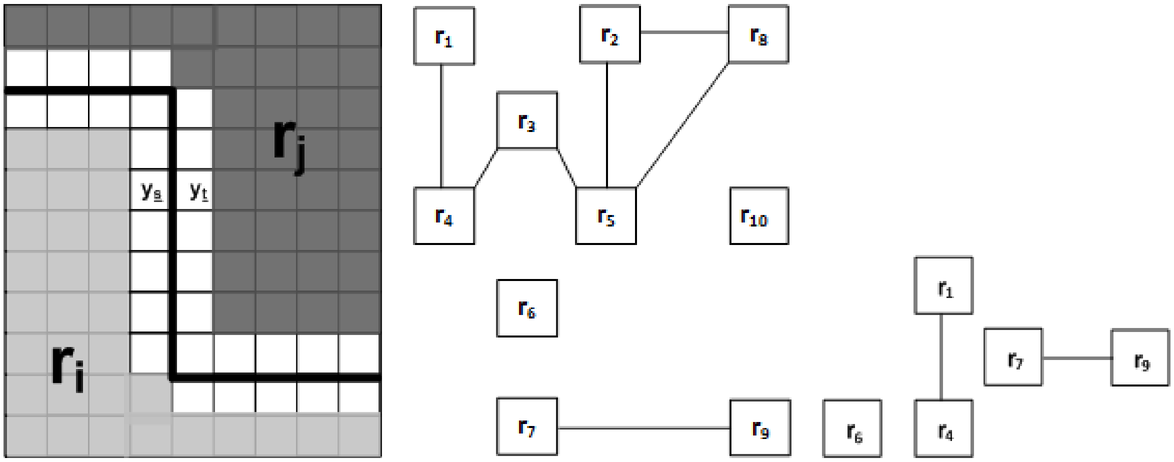


Fig. 3.1: (a) Example of unsupervised classification given polygons with labeling constraints.(b)Example of final classification, labeling of regions. $r = \{r_1, \ldots, r_{10}\}$, $L(P_1) : \{r_1, r_2, r_3\} \to \{l_1, l_3, l_4\}$, $L(P_2) : \{r_4, r_5, r_6, r_7\} \to \{l_1, l_2, l_3, l_4\}$, $L(P_3) : \{r_8, r_9, r_{10}\} \to \{l_1, l_2, l_3\}$

where the *a priori* term $p(Z = z)$ in MRF can be modeled as a Gibbs distribution Eq. 3.3 and the feature model $p(X = x|Z = z)$ as an example can be Gaussian mixture model (GMM). If the Gibbs and Gaussian probability density functions are substituted into



(a) (b) (c)

Fig. 3.2: Boundary, graph, cliques. (a) The boundary is comprised of all white pixels connecting two regions $r_i, r_j$ based on first-order neighborhood system [88]. The bold black line is an edge separating two regions. White pixels along the edge are part of regions. (b) Graphical representation of Fig. 3.1(a). (c) Example of first-order cliques of $\mathsf{C} = [\mathsf{c} \mid \mathsf{c} \subset N_r]$.

Eq. 3.5, 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 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 labeling can be accomplished by solving the following optimization problem:

$$\arg \min_{\{z_r | r \in S\}} E_{total} \tag{3.6}$$

where,

$$E_{total} = \alpha E_f + \beta E_p \tag{3.7}$$

is the total energy. $\alpha$ and $\beta$ are weighting parameters described in Section 3.3.

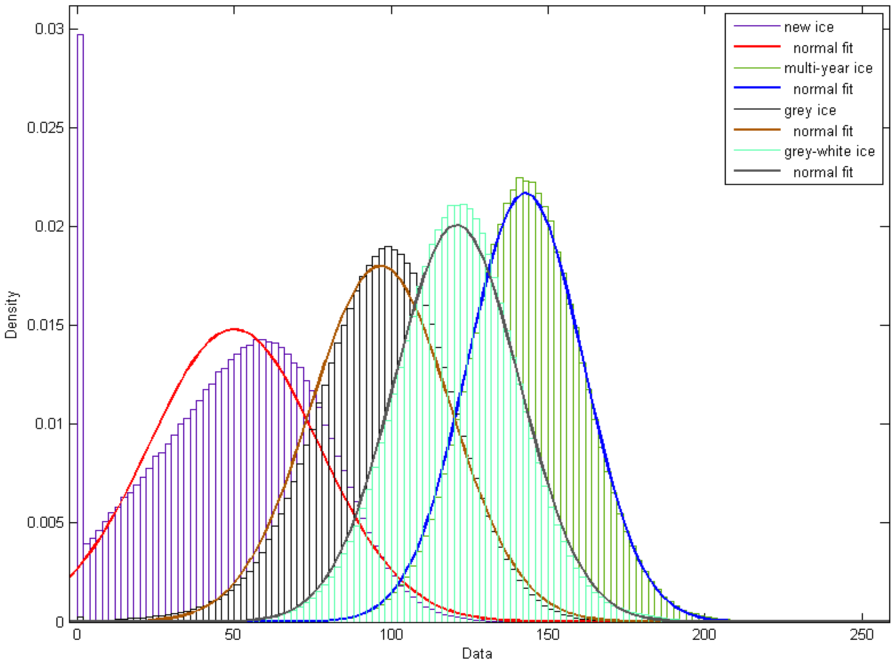
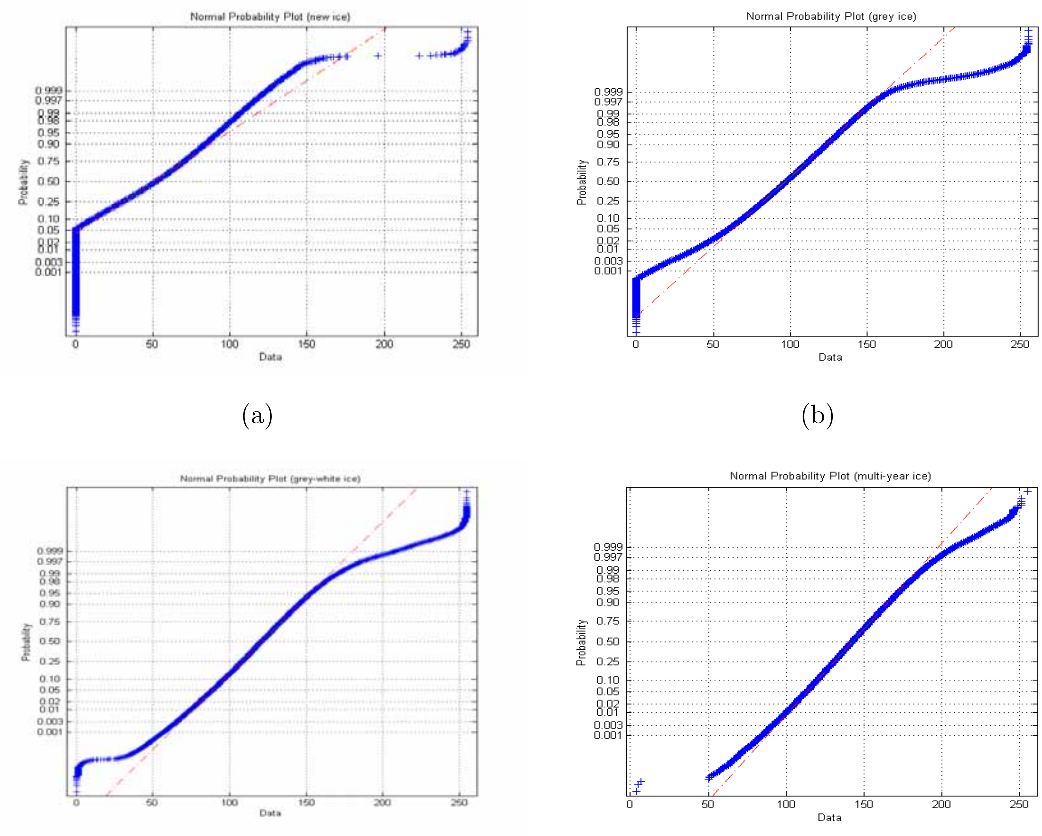


Fig. 3.3: The overlayed density plots of four classes from SAR imagery. High overlap is seen.

### Feature model $E_f$

SAR and images of nature are the main data sources used in the thesis. The statistical nature of SAR images indicates that the amplitude of the scattered signal is gamma distributed, however, in-house testing and published research [74] indicate that modeling classes in the feature space as Gaussian produces acceptable results. Here are the examples from used data. The overlayed density plots of four classes from SAR imagery is depicted in Fig. 3.3 and the corresponding normal probability plots for each class are shown in Figs. 3.4(a),(b),(c) and (d). Although, bell-shaped density plots are seen in Fig. 3.3 deviations present in the normal fit plots. The closer analysis can be performed from visual analysis of the linearity in normal probability plots (Figs. 3.4(a),(b),(c) and (d)). Although linear relationship is observed there are deviations in all classes which is due to the variability in class statistics.



(a) (b)

(c) (d)

Fig. 3.4: Normality test of SAR image classes. (a) Normal probability plot of new ice. (b) Normal probability plot of grey ice. (c) Normal probability plot of grey-white ice. (d) Normal probability plot of multi-year ice.

The linearity seen in normal probability plots of an example natural image is higher for homogeneous classes such as the background (Fig. 3.6(a)) and grass (Fig. 3.6(b)), but less in class containing the bird (Fig. 3.6(c)). The same observation can be made from normal fit in Fig. 3.5. Although challenging classes exists, in average, some level of linearity in normal probability plots of classes is present and the mixture of Gaussian distributions is the generic model that can describe the data used. Thus, the energy of a single region with

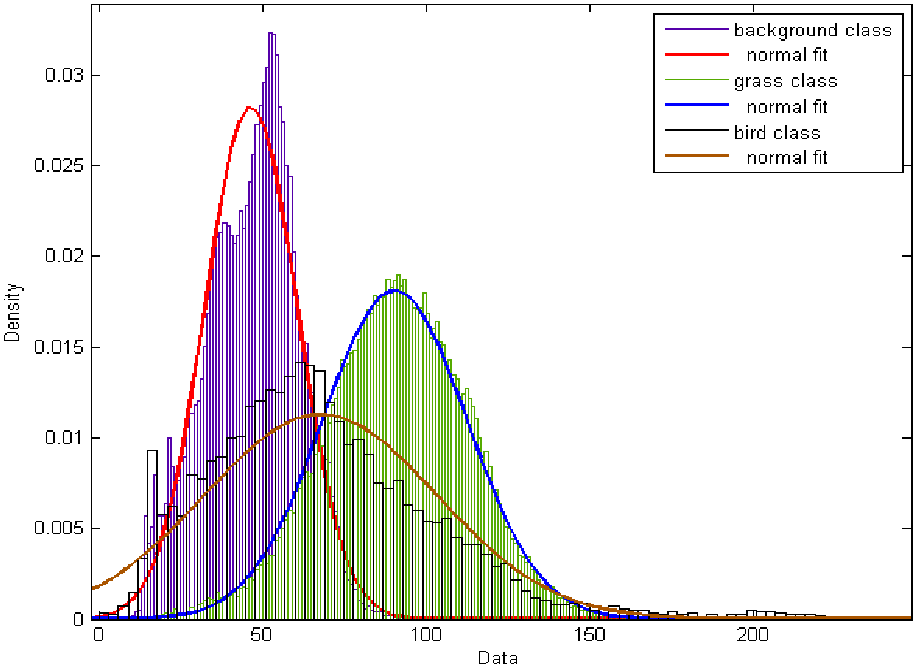
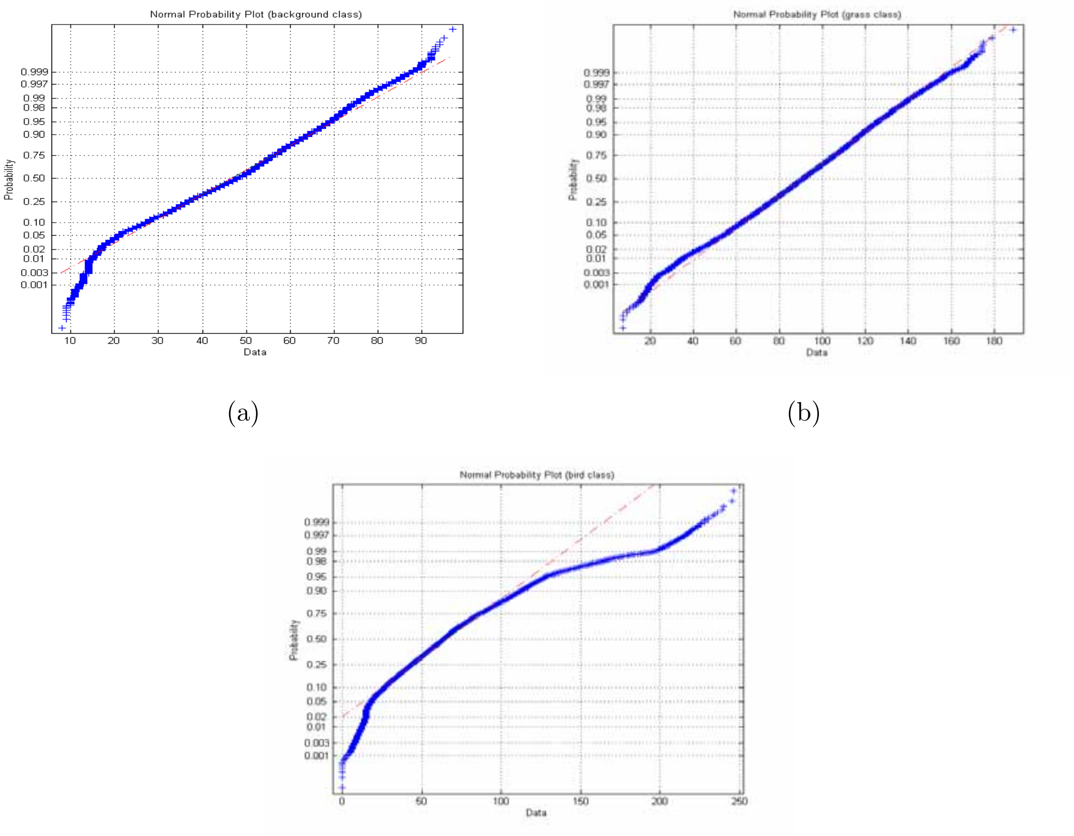


Fig. 3.5: The overlayed density plots of three classes from image of nature. High overlap is seen.

respect to a class can be derived as:

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



(a) (b)

(c)

Fig. 3.6: Normality test of nature image classes. (a) Normal probability plot of background class. (b) Normal probability plot of grass class. (c) Normal probability plot of bird class.

$$= k_{r_i}\frac{1}{2}\ln 2\pi\sigma_{l_{r_i}}^2 + \frac{k_{r_i}}{2\sigma_{l_{r_i}}^2}(u_{r_i} - u_{l_{r_i}})^2 +$$

$$+ \frac{(k_{r_i} - 1)\sigma_{r_i}^2}{2\sigma_{l_{r_i}}^2}$$

for very large $k_{r_i}$, $k_{r_i} - 1 \simeq k_{r_i}$ and the expressioncan be simplified as

$$E(r_i) = \frac{k_{r_i}}{2\sigma_{l_{r_i}}^2}\left[\ln 2\pi\sigma_{l_{r_i}}^4 + (u_{r_i} - u_{l_{r_i}})^2 + \sigma_{r_i}^2\right] \tag{3.8}$$

where parameters $u_{r_i}$,$\sigma_{r_i}^2$,$k_{r_i}$ are the mean, covariance and the number of pixels of region $r_i$. In any realization of $Z$ all $r_i$ are assigned a label $l_{r_i}$. The term $u_{l_{r_i}}$, $\sigma_{l_{r_i}}^2$ are the estimates of all $r_i$ with the same label $l_{r_i}$. Equation (3.8) 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:

$$E_f = \sum_{i=1}^{n_r} E(r_i) \tag{3.9}$$

### A priori term $E_p$

The regional representation of the standard multilevel logistic model MLL [88] can be defined as the energy of the Gibbs distribution with zero single node clique energy and pairwise clique energy given as:

$$E_p = \sum_{i=1}^{n_r} \left[ \sum_{r_j \in N_{r_i}} \delta(l_{r_i}, l_{r_j}) \right] \tag{3.10}$$

where $l_{r_i}$ and $l_{r_j}$ are the labels assigned to $r_i, r_j$ respectively. Equation 3.10 favors the configurations of neighboring regions having the same label. Such an *a priori* model is not suitable to model polygon interactions in the bootstrap approach. Polygons can be generated to divide a large region with the same class into smaller regions or to separate regions with different classes, therefore two neighboring regions are not necessarily always of the same class. To accommodate such an interaction, an edge penalty term has been incorporated into Eq. 3.10.

$$E_p = \sum_{i=1}^{n_r} \left[ \sum_{r_j \in N_{r_i}} g(\nabla_{r_i,r_j}) \delta(l_{r_i}, l_{r_j}) \right] \tag{3.11}$$

where,

$$g(\nabla_{r_i,r_j}) = 1 - \nabla_{r_i,r_j} \tag{3.12}$$

is the edge penalty term with the edge strength $\nabla_{r_i,r_j}$ normalized to range $[0 \ldots 1]$. The edge strength is calculated by taking the absolute value of first order difference [5] between the neighboring pixels in the boundary shown in Fig 3.2:

$$\nabla_{r_i,r_j} = \sum_{s,t \in \partial} |y_s - y_t| \tag{3.13}$$

Edge penalty increases if the edge strength is small; thereby, Equation 3.11 dictates that neighboring regions having low edge strength are likely to belong to the same class.

Implementation of edge penalty similar to GIEP [2] is not feasible for processing large scenes since it requires recalculation of edge penalty at every iteration step since the varying parameter is embedded in the power of exponential function which requires all the neighboring boundary pixels(Eq.18 in [2]). In this research the purpose is to minimize the pixel-based processing, therefore, the edge penalty is not growing, calculated only once and stored for every pairwise cliques by basic negative operation in Eq. 3.12 which increases edge penalty if the edge strength is small. Nevertheless, the variability and nonstationarity is transformed to the weight $\alpha$ of feature model as in V-MLL [90]. However, the V-MLL model does not use edge penalty as a spatial term, therefore the proposed model is different from V-MLL [90] and GIEP [2].

## 3.3 Implementation Scheme

A standard approach is used to find an optimal solution. Given Eq. 3.9 and Eq. 3.11 the combination of simulated annealing (SA) [122] and Metropolis sampling [123] have been applied using a common temperature schedule [124].

There are four parameters to be estimated: $\alpha$, $\beta$, $u_{l_{r_i}}$ and $\sigma^2_{l_{r_i}}$ $\forall i$ where $r_i$ has been assigned the same label. As an unsupervised algorithm, the EM [125, 126] algorithm can be used for estimating the class mean and covariance over a full scene. EM is suitable for maximum likelihood estimation of feature parameters of incomplete data [125] and the convergence of the EM algorithm is known [127].

Here, the parameter $\beta$ is set to one and the parameter $\alpha$ is estimated. In conventional MLL models [88], $\alpha$ 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 [90] to improve performance. As such, the parameter $\alpha$ can be a function of iteration $\theta$:

$$\alpha(\theta) = c_1\, 0.9^{\theta} + c_2 \tag{3.14}$$

where $c_1, c_2$ are constants equal to 0.1 [90]. Equation 3.14 monotonically decreases the parameter $\alpha$ with each iteration.

## 3.4 Algorithm Flow and Computational Efficiency

The algorithm is presented in seven modules as shown in Fig. 3.7 with each step defined below. While dealing with the typical large images the computational demands should be addressed for operational requirements.

I The polygon data, the unsupervised classification result and the original image are loaded into memory. Therefore, the memory of the computing platform must be sufficient for handling at least two full scenes. For SAR scenes the average size is 800 MBytes.

II The neighborhood is obtained using the spatial relationship of regions as specified in Fig. 3.2(b) 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 image in raster format with the computational complexity related to number of the image pixels $O(MN)$. Hence, in the proposed technique, the edge penalty calculation in Eq. 3.12 is performed once and stored as the regional edge penalty used for all iterations.

III The associated mean $u_{r_i}$, variance ${\sigma_{r_i}}^2$ and number of samples $k_{r_i}$ are stored for each region. These three values are sufficient to calculate the energy for $r_i$ (Eq. 3.8). The computation of feature statistics is the last intensive processing step and has the complexity $O(MN)$.

IV The labels are randomly assigned to regions as per *Definition 1* to obtain the initial labeling realization $Z$.

V Using the $u_{l_{r_i}}$ and $\sigma^2_{l_{r_i}}$ from the previous step, Eq. 3.7 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 (Eq. 3.8). 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. 3.7 with SA and Metropolis sampling the labeling refinement is accepted or rejected in the E-step using:

$$\rho > \exp\left[\frac{E_1 - E_2}{T}\right] \tag{3.15}$$

where $\rho$ 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 [124].

VI Using the labeled image from the previous step, $u_{l_{r_i}}$ and $\sigma^2_{l_{r_i}}$ $\forall i$ are estimated by iteratively building the statistics from unions of regions $r_{ij} = r_i \cup r_j$ having the same label. The formulation to achieve this task can be given as in [121, p. 119]:

$$\begin{aligned} k_{r_{ij}} &= k_{r_i} + k_{r_j} \\ u_{r_{ij}} &= \frac{u_{r_i} k_{r_i} + u_{r_j} k_j}{k_{ij}} \\ \sigma^2_{r_{ij}} &= \frac{\sigma_{r_i} k_{r_i} + \sigma_{r_j} k_{r_j}}{k_{r_{ij}}} + \\ &+ \frac{(u_{r_i} - u_{r_j})^2 k_{r_i} k_{r_j}}{k^2_{r_{ij}}} \end{aligned} \tag{3.16}$$

where $k_{r_{ij}}$, $u_{r_{ij}}$ and $\Sigma_{r_{ij}}$ 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.

VII After energy oscillation at initial iterations the system is cooled down using a temperature schedule [124] to settle at global minima. The algorithm always converges well before reaching the set 100 iterations.

## 3.5 Evaluating the Methodology

The polygon generation, unsupervised classification of individual polygons and the labeling of regions are the components contributing to the performance of classification with bootstrap approach. The reference images are produced at different stages to better evaluate the performance of individual components and of overall system. Thus, the evaluation is composed of following stages of complexity:

1. At first stage the evaluation is performed using a dataset [128] with perfect unsupervised classification and polygon boundaries derived from reference images (Chapter 4).

2. At the next stage, the evaluation is performed with operational reference images (Section 5.1.2) where the expert evaluator trusts the unsupervised classification and provides the labels for the regions. The second stage is more difficult since the polygon boundaries are not ideal (Chapter 5).
3. The third stage is a challenging evaluation since the evaluator does not trust the unsupervised classification and provides a reference image (Section 6.4.2) by manually labeling the pixels. Unsupervised classification and polygon boundaries are both not ideal in this case (Chapter 6).

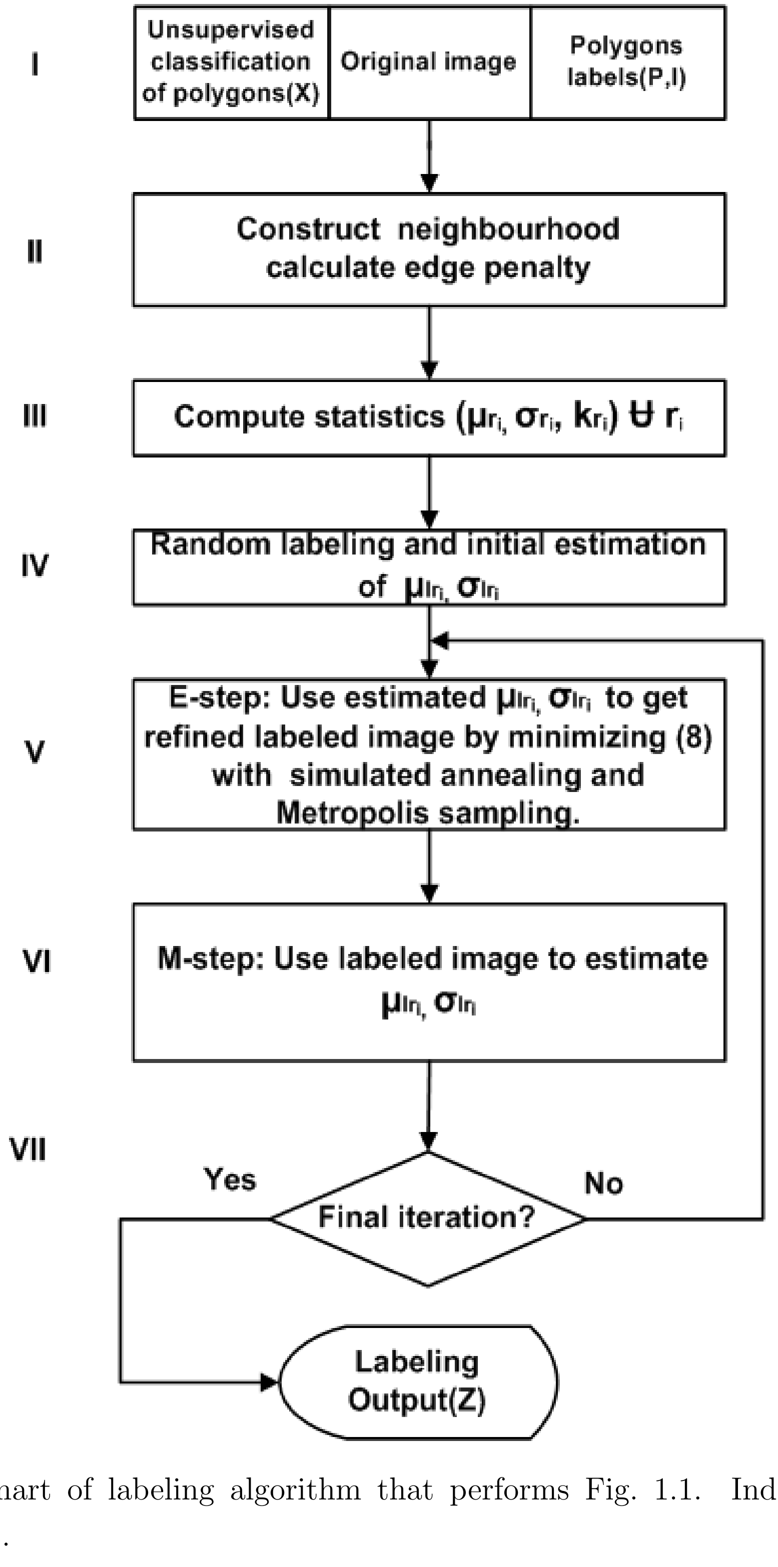


Fig. 3.7: Flowchart of labeling algorithm that performs Fig. 1.1. Indicated steps are discussed in text.

# Chapter 4

# Nature Image Classification

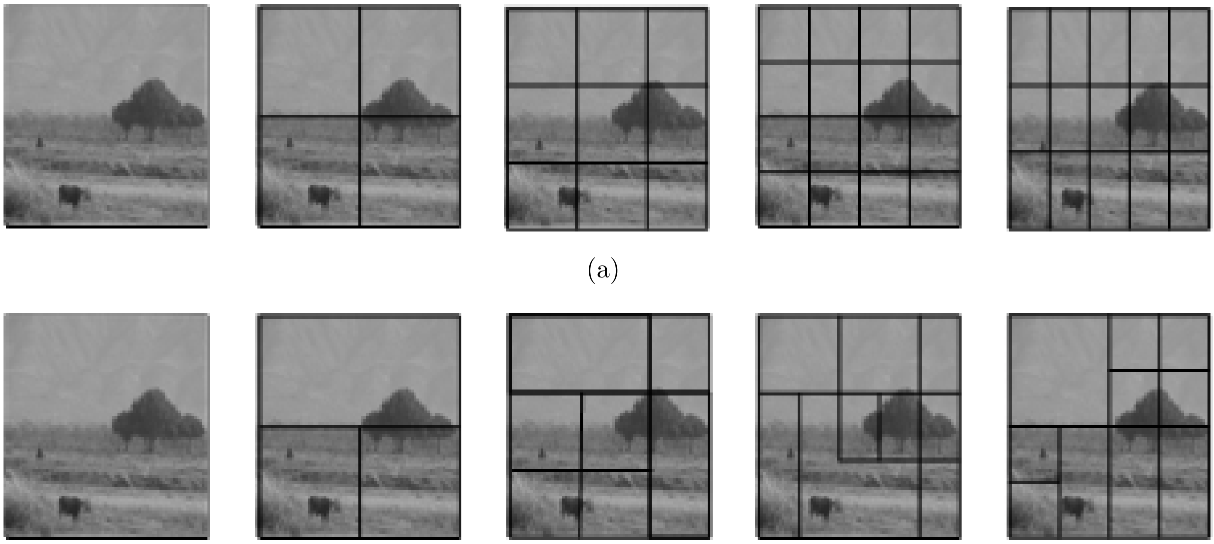
The concepts Chapter 3 outlined have been applied to scenes of nature and artificial images. The framework has been presented for generation of polygons. Unsupervised classification of polygons is assumed ideal in this scenario; hence, the focus is on performance evaluation of image labeling component of bootstrap approach. Results indicate the importance of *a priori* term, robustness of technique to image distortions and noise.

## 4.1 Polygon Generation and Ideal Unsupervised Classification

### 4.1.1 Polygon generation

Two methods of polygon generation are introduced:

- A fast way of generating polygons is to automatically partition the image into rectangles. The purpose is to avoid generation of too small polygons by generating polygons that are the same or nearly the same size. The examples of partitioning is shown in Fig. 4.1(a). The requirement for selecting the number of polygons is to have multiple repeating classes in different polygons. The existence of same classes in different polygons (Fig. 4.2(a)) is what makes the automatic labeling technique successful. Continuity between polygons is another important factor affecting the performance.



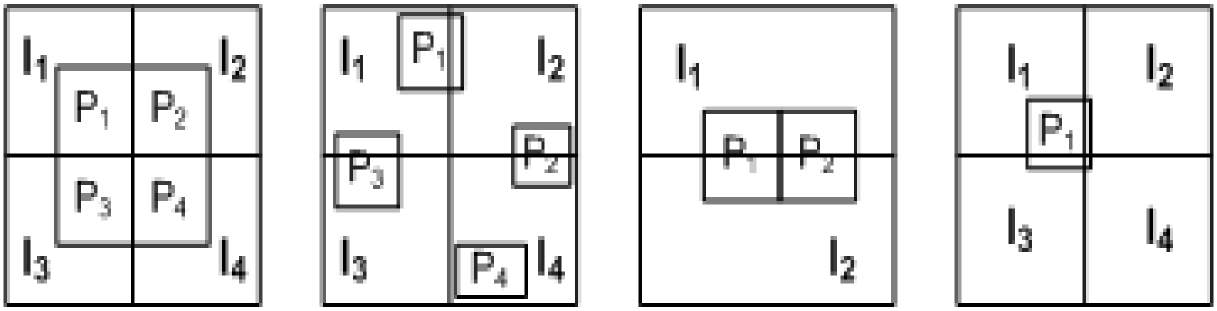
(a)

(b)

Fig. 4.1: Examples of polygon generation. (a) The image is partitioned into rectangles. Four examples are given with different number of polygons. (b) The image is partitioned into polygons but some of them are combined to form a larger polygon. Four examples are given with different polygon combinations.

Isolated non-contiguous polygons as shown in Fig. 4.2(b) may deteriorate the accuracy. Unconstrained polygons in Fig. 4.2(c) may result in absence of unique labeling solution. In this case, the labeling solution helps combine the regions to classes, but without assigning a qualitative label. The polygons should minimally contain the few pixel classes as it is the case in Fig. 4.2(d). This type of polygon generation has been used for natural scene labeling in Section 4.2.

- Another way is to separate the image into polygons where the user can combine objects/regions to construct the larger polygons as shown in examples in Fig. 4.1(b). The requirements discussed in previous way are valid here as well. This type of approach has been used for artificial scene labeling in Section 4.3.



(a)

Fig. 4.2: Limitations in polygon generation. The polygon generation examples presented in the figure must be avoided. (a) The non-existence of same classes in different polygons. (b) Isolated non-contiguous polygons. (c) Unconstrained polygons. Both polygons have same set of classes no unique solution. (d) Polygon having few pixel classes ($l_4$).

### 4.1.2 Unsupervised classification

The unsupervised classification part of the bootstrap approach (Section 2.6.4) is not performed in this chapter. Unsupervised classification results have been obtained from reference images provided by a third party. The purpose is to focus on automatic labeling performance by having ideal polygons and unsupervised classification results.

### 4.1.3 Performance metrics

As a performance metric, accuracy is calculated as a percentage of correctly classified pixels and regions, and the kappa coefficient ($\kappa$) is used as a means of classification agreement [129]:

$$\kappa = \frac{P(A) - P(E)}{1 - P(E)} \tag{4.1}$$

where $P(A)$ is the probability the model values are equal to the actual value and $P(E)$ is the expected probability by chance.

A value of one for kappa coefficient means a statistically perfect classification while a 0 means that all values are randomly classified. Generally, a $\kappa \geq 0.7$ indicates accurate statistical modeling. The accuracy and the kappa metric are calculated based on pixel-level or region-level accuracy. Pixel accuracy is the percentage of pixels properly classified. As

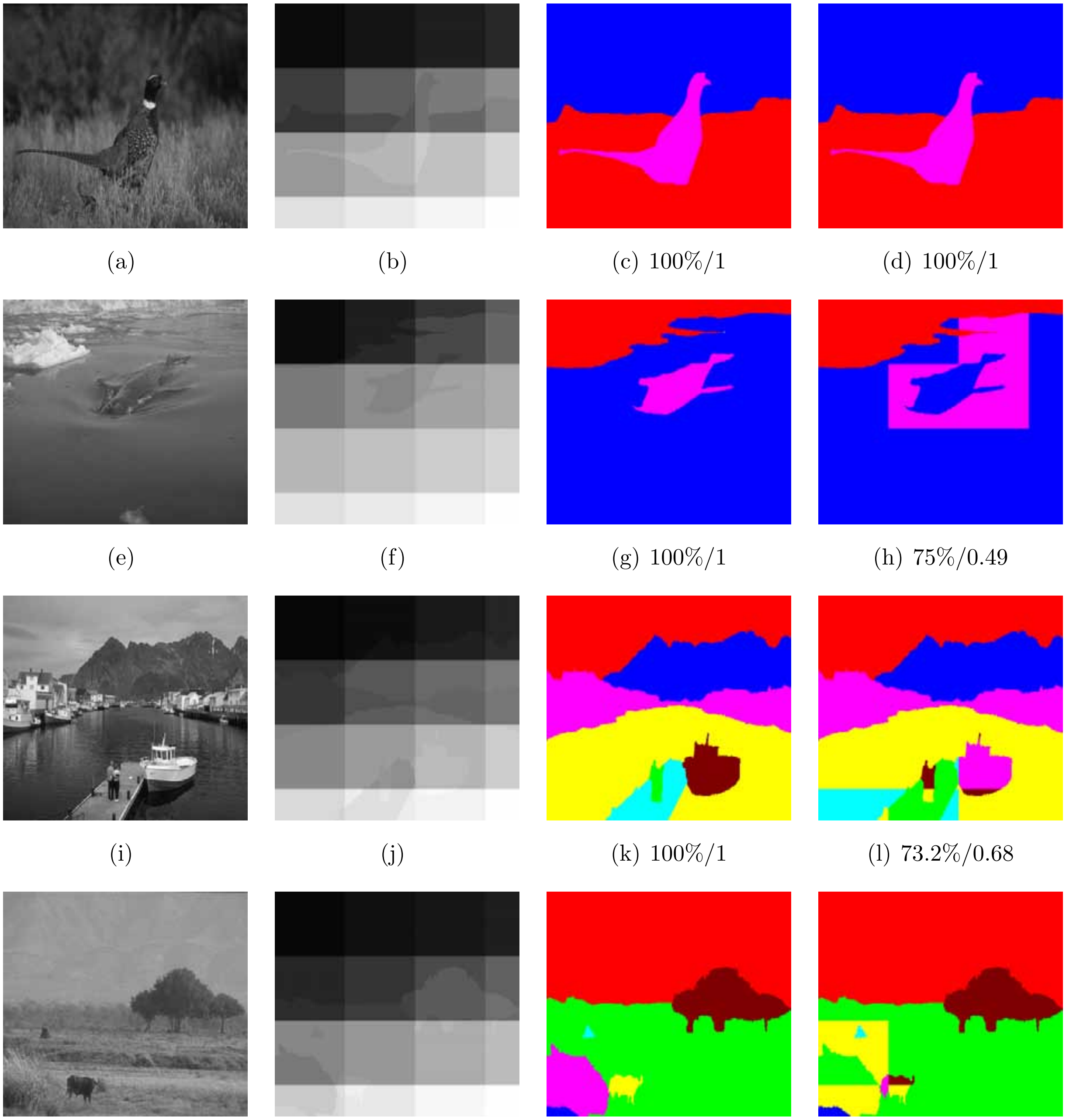
Table 4.1: Example of ground reference information table for calculating P(A) and P(E) values.

| class | classA | classB | Total |
|---|---|---|---|
| classA | 2 | 5 | 7 |
| classB | 10 | 83 | 93 |
| Total | 12 | 88 | 100 |

a result, the pixel accuracy involves the size of the regions, whereas, the region accuracy is irrespective of the region size. An example of how P(A), P(E) and $\kappa$ is calculated can be presented using the truth table 4.1 as: $P(A) = \frac{2+83}{100} = 0.85$, $P(E) = [(7/100) * (12/100)] + [(93/100) * (88/100)] = 0.008 + 0.81 = 0.82$ and $\kappa = \frac{0.85-0.82}{1-0.82} = 0.16$.

## 4.2 Natural Scene Labeling

Natural scenes used in this chapter are part of the Berkley segmentation dataset [128]. The first column of Fig.4.3 depicts the original images of various scenes. The second column shows images partitioned into polygons and unsupervised classification results are obtained as described in Section 4.1. The third column depicts the labeling results which indicate perfect accuracy. The setup in Figs. 4.3(a) and (e) is simpler with three classes well mixed in 16 polygons. The scenario in Fig. 4.3(i) and (m) is challenging with seven classes distributed in only 16 polygons. Nevertheless, the bootstrap approach yields the perfect labeling for Fig.4.3(k) and (o). Even though the implementation of GMM with the bootstrap approach is the novel technique by itself, it is presented in this thesis to emphasize the importance of spatial interactions. The fourth column depicts the result of the bootstrap approach with GMM and the lower accuracy obtained for images in Fig. 4.3(h), Fig. 4.3(l) and (p) where the class statistics alone are not sufficient to accurately label the images. For example in Fig. 4.3(h) the fish has a tone close to darker water and therefore they have been combined into the same class. The brighter water around the fish has been identified as a separate class; however, the proposed method in Fig. 4.3(g) successfully identified fish and



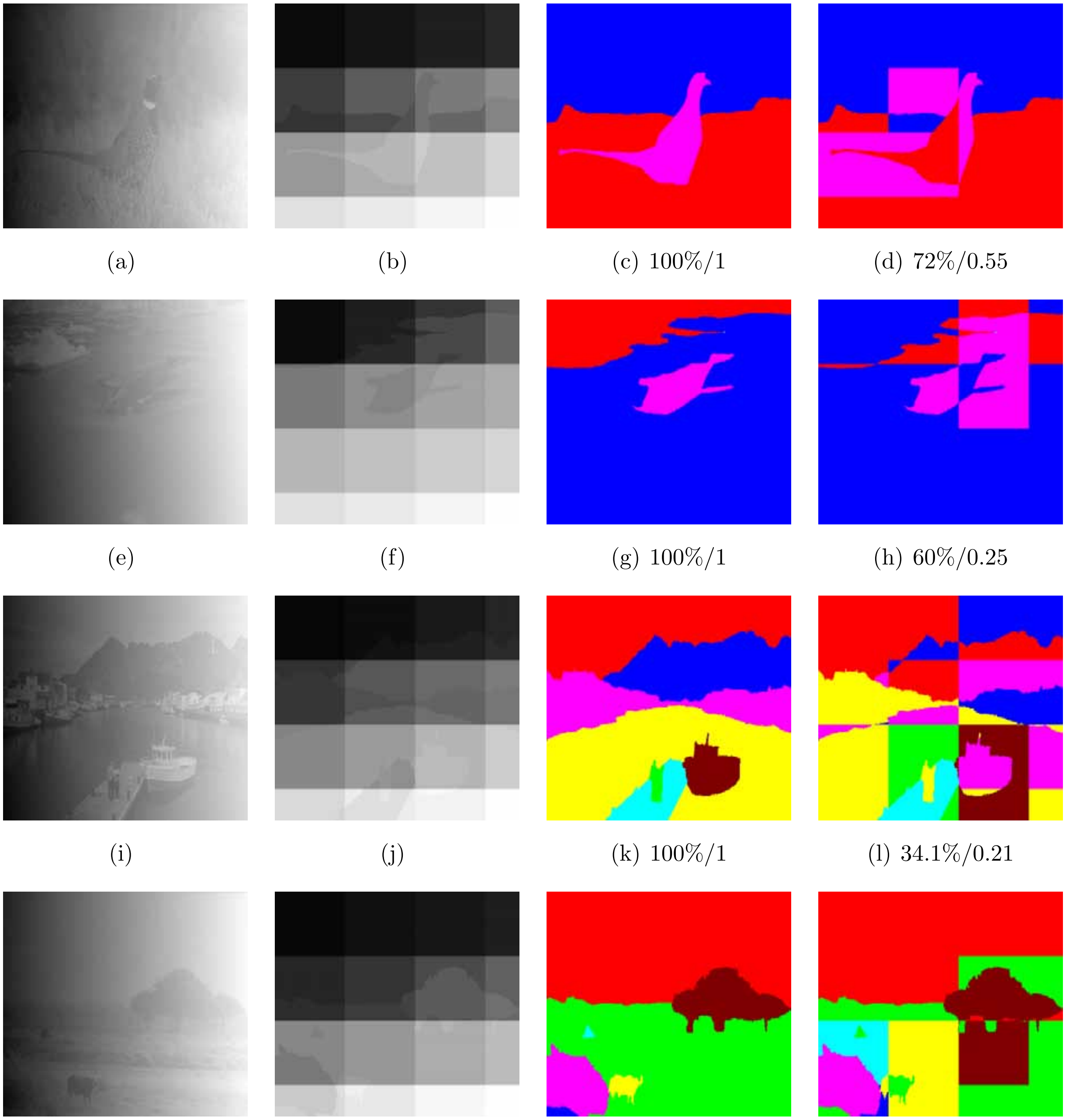
(a) (b) (c) 100%/1 (d) 100%/1

(e) (f) (g) 100%/1 (h) 75%/0.49

(i) (j) (k) 100%/1 (l) 73.2%/0.68

(m) (n) (o) 100%/1 (p) 79.4%/0.73

Fig. 4.3: Natural image labeling results. (a-e-i-m) Original images. (b-f-j-n) Images separated into 16 polygons. (c-g-k-o) Labeled images with proposed method(*accuracy*/*kappa* averaged over 10 runs). (d-h-l-p) Labeled images with proposed method but GMM (*accuracy*/*kappa* averaged over 10 runs).


(a) (b) (c) 100%/1 (d) 72%/0.55

(e) (f) (g) 100%/1 (h) 60%/0.25

(i) (j) (k) 100%/1 (l) 34.1%/0.21

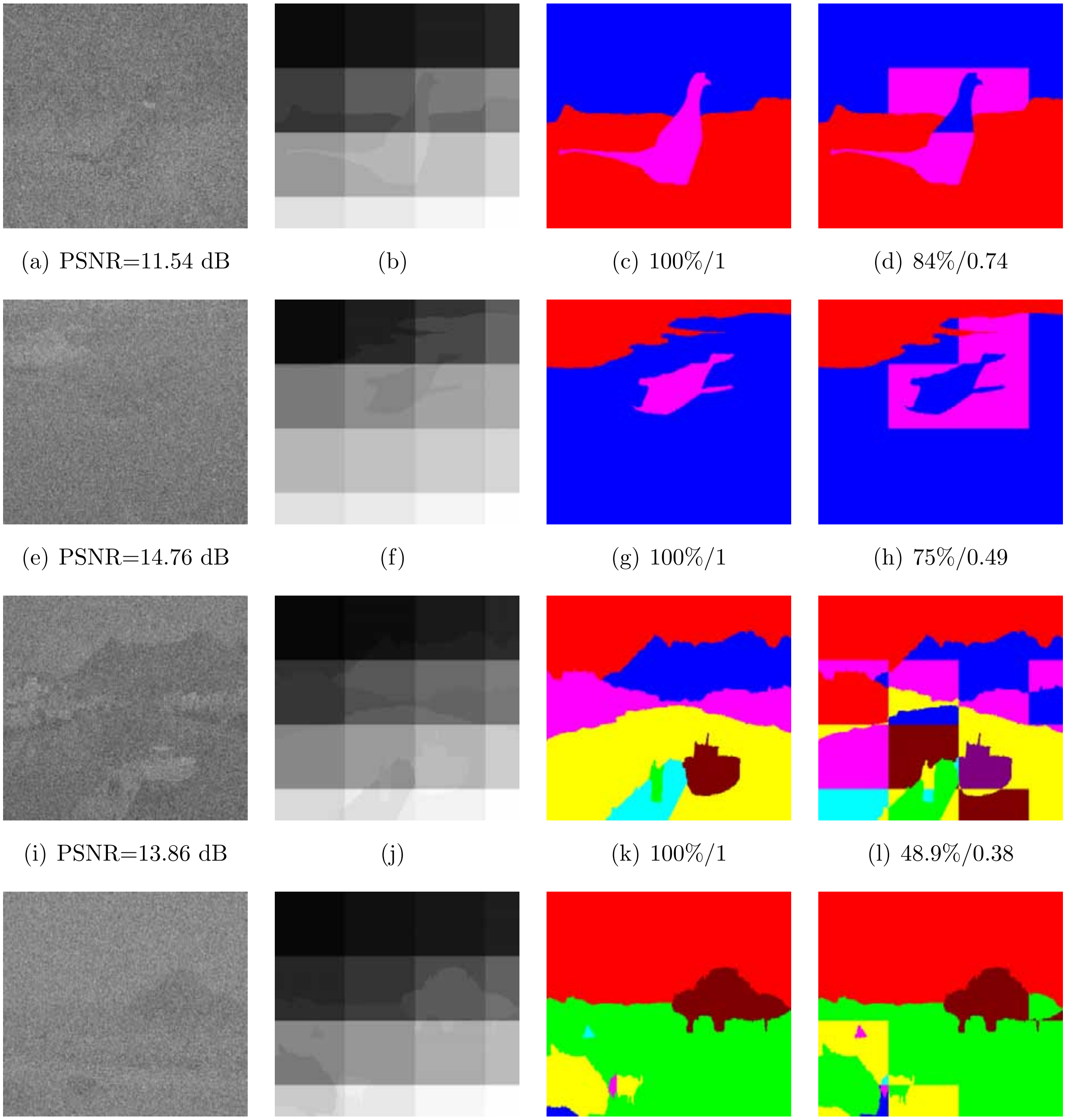
(m) (n) (o) 88.6%/0.85 (p) 58.8%/0.46

Fig. 4.4: Natural image labeling results with illumination distortion. (a-e-i-m) Original images with illumination distortion. (b-f-j-n) Images separated into 16 polygons. (c-g-k-o) Labeled images with proposed method (*accuracy*/*kappa* averaged over 10 runs). (d-h-l-p) Labeled images with proposed method but GMM (*accuracy*/*kappa* averaged over 10 runs).

water as single separate class, similarly, the boat and the white city buildings in Fig. 4.3(l) are merged, resulting in misclassification with proposed method but GMM, whereas the buildings in Fig. 4.3(k) have been combined into the same class due to proximity, but the boat remained as separate class resulting in perfect labeling with proposed method. The same outcome was obtained for the tree and cow in Fig. 4.3(p) and (l).

Illumination distortion has been added to natural scenes in Fig. 4.4. The illumination distortion is a constant added to every column of an image and is gradually increasing from the left to right direction. In this case the constant is 2 and incremented by 1 as the window is moved from left to right. The scenario can be clearly seen in the first column of Fig. 4.4, where the variability of class statistics within the image is significant. The third column in Fig. 4.4 presents the results with the bootstrap approach, where three images have been perfectly labeled and last image in Fig. 4.4(o) has only two misclassified regions. Similarly, to emphasize the importance of the prior term, the results in the fourth column of Fig. 4.4 have been obtained using the same bootstrap approach, but with GMM. The result indicate the low accuracy and high misclassification rate when the prior term is not considered. The regular scenario in Fig. 4.3 yields low performance with GMM; therefore, high variability in class statistics due to illumination distortion has further reduced performance of labeling with proposed method but GMM.

The experiment has been performed with natural images by adding Gaussian noise. The original images with corresponding peak signal-to-noise ratio (PSNR) values are depicted in Fig. 4.5 first column. The same polygon generation and ideal unsupervised segmentation, used in previous experiments, depicted in second column of Fig. 4.5. As shown in Fig. 4.5 third column, the performance of the proposed method remained the highest. The resultant images in Fig. 4.5 have been averaged over ten executions and presented values correspond to the lowest PSNR. Furthermore, the same experiment has been performed with different variance of noise and the results are provided in Table. 4.2 for all four images. The consistent performance of the proposed method proves the robustness of the method against additive noise. One notable difference has been observed in the results of proposed method with GMM. Compared to previous results in Fig. 4.3 and Fig. 4.4 the performance of proposed method with GMM has increased with an addition of Gaussian noise. The increase in performance is expected since the class characteristics approaches the ideal Gaussian distribution due to noise as can be seen from the normal distribution plot in Fig. 4.6, where the background and grass class have exhibited near perfect linear behavior.



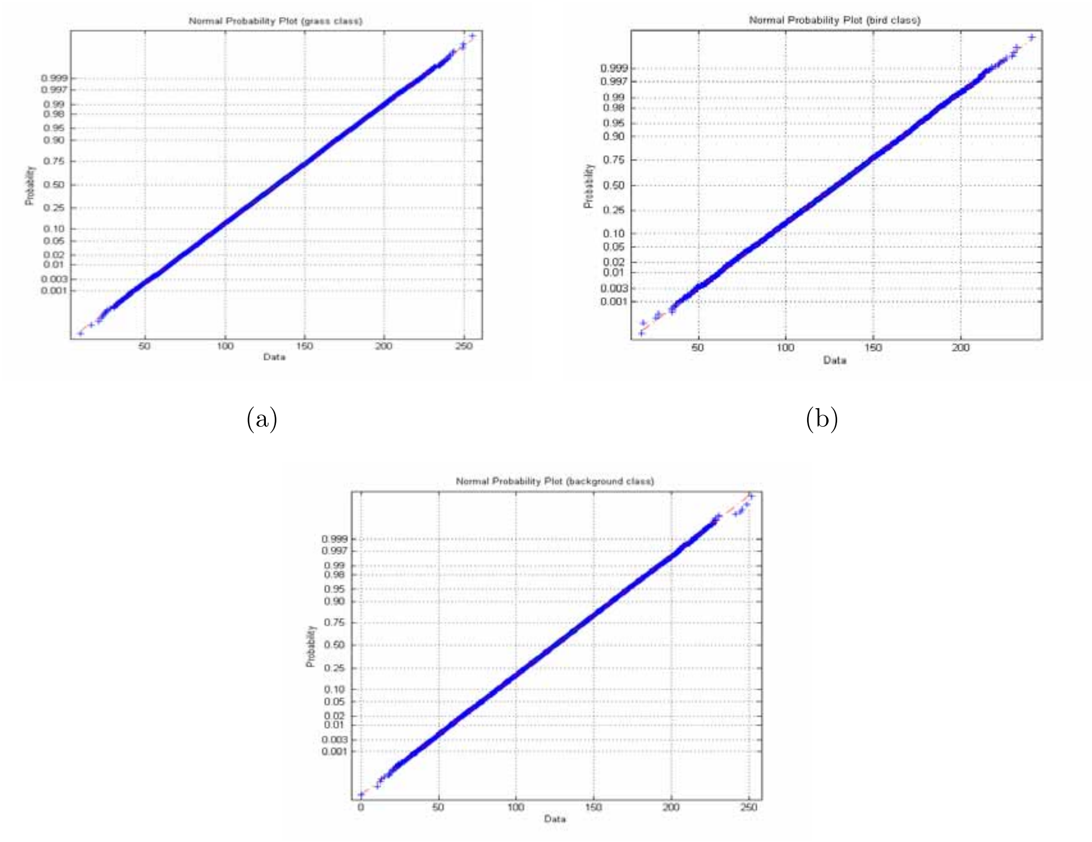
(a) PSNR=11.54 dB (b) (c) 100%/1 (d) 84%/0.74

(e) PSNR=14.76 dB (f) (g) 100%/1 (h) 75%/0.49

(i) PSNR=13.86 dB (j) (k) 100%/1 (l) 48.9%/0.38

(m) PSNR=16.56 dB (n) (o) 88.2%/0.84 (p) 64.7%/0.54

Fig. 4.5: Natural image labeling results with additive Gaussian noise. (a-e-i-m) Original images with added Gaussian noise. (b-f-j-n) Images separated into 16 polygons. (c-g-k-o) Labeled images with proposed method (*accuracy*/*kappa* averaged over 10 runs). (d-h-l-p) Labeled images with proposed method but GMM (*accuracy*/*kappa* averaged over 10 runs).



(a) (b)

(c)

Fig. 4.6: Normality test of nature image classes with additive Gaussian noise (PSNR=11.54 dB). (a) Normal probability plot of background class. (b) Normal probability plot of grass class. (c) Normal probability plot of bird class.

## 4.3 Artificial Image Labeling

The experiment has been performed with artificial reference image having ideal Gaussian distributed classes. The artificial reference image in Fig. 4.7(a) is produced by randomly selecting rectangles from one of the classes. The image is divided into 64 equal sized regions

Table 4.2: Summary of classification results (*accuracy*/*kappa* averaged over 10 runs) for four images in Fig. 4.5, respectively, with additive noise of different variance.

| $\delta_n^2 x10^2$ | PSNR (dB) | Proposed method | Proposed method with GMM |
|---|---|---|---|
| 10 | 13.04 | 100%/1 | 100%/1 |
| 20 | 12.07 | 100%/1 | 100%/1 |
| 30 | 11.90 | 100%/1 | 100%/1 |
| 90 | 11.81 | 100%/1 | 84%/0.74 |
| 100 | 11.54 | 100%/1 | 84%/0.74 |

| $\delta_n^2 x10^2$ | PSNR (dB) | Proposed method | Proposed method with GMM |
|---|---|---|---|
| 10 | 17.96 | 100%/1 | 75%/0.49 |
| 20 | 16.12 | 100%/1 | 75%/0.49 |
| 30 | 15.59 | 100%/1 | 75%/0.49 |
| 90 | 14.95 | 100%/1 | 66.7%/0.32 |
| 100 | 14.76 | 100%/1 | 75%/0.49 |

| $\delta_n^2 X10^2$ | PSNR (dB) | Proposed method | Proposed method with GMM |
|---|---|---|---|
| 10 | 16.59 | 100%/1 | 53.7%/0.44 |
| 20 | 14.74 | 100%/1 | 48.8%/0.38 |
| 30 | 14.18 | 100%/1 | 56.1%/0.47 |
| 90 | 13.40 | 100%/1 | 56.1%/0.47 |
| 100 | 13.86 | 100%/1 | 48.9%/0.38 |

| $\delta_n^2 x10^2$ | PSNR (dB) | Proposed method | Proposed method with GMM |
|---|---|---|---|
| 10 | 18.47 | 88.2%/0.84 | 82.3%/0.77 |
| 20 | 17.89 | 88.2%/0.84 | 76.4%/0.69 |
| 30 | 17.59 | 88.2%/0.84 | 76.4%/0.69 |
| 90 | 16.81 | 88.2%/0.84 | 64.7%/0.54 |
| 100 | 16.56 | 88.2%/0.84 | 64.7%/0.54 |

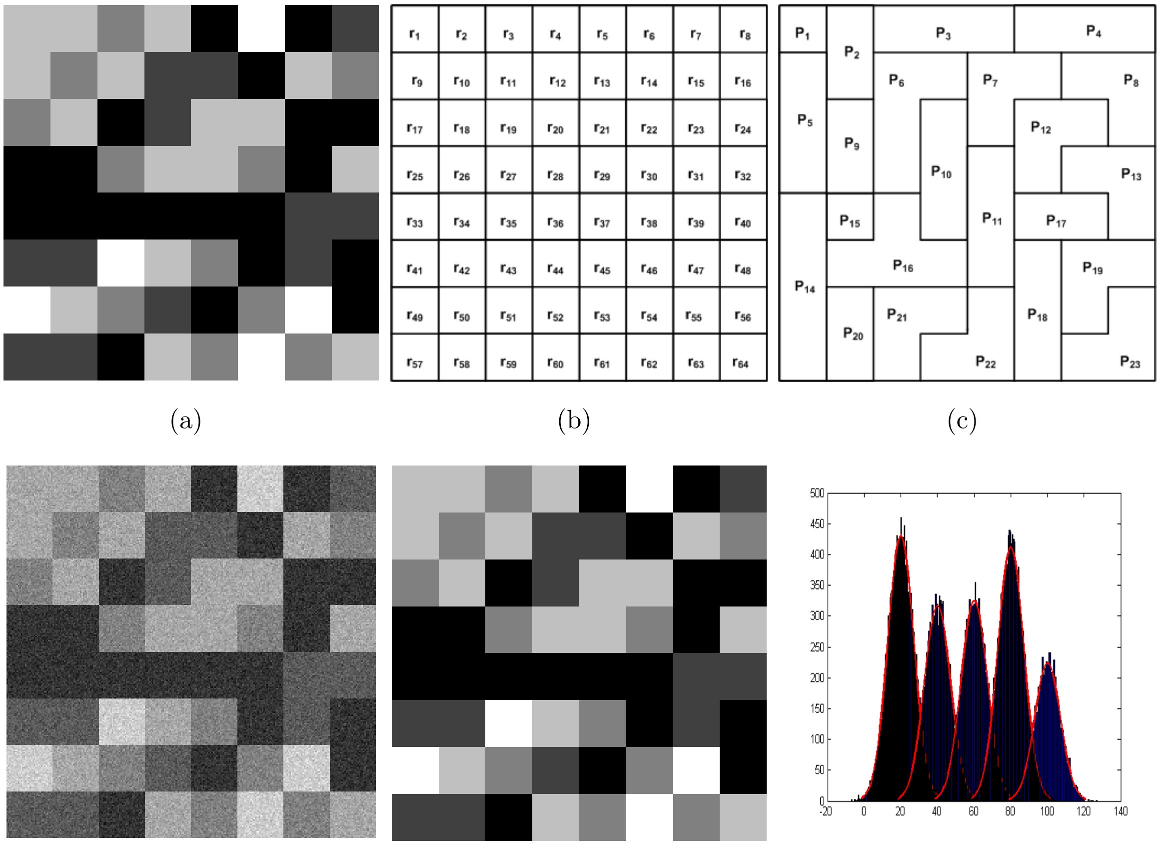
Table 4.3: Summary of labeling results (*accuracy*/*kappa* averaged over 10 runs) for artificial image (Fig. 4.7) with additive noise of different variance.

| $\delta_n^2$ | PSNR (dB) | Proposed method | Proposed method with GMM |
|---|---|---|---|
| 10 | 30.8840 | 100%/1 | 100%/1 |
| 20 | 28.2422 | 100%/1 | 100%/1 |
| 30 | 27.1827 | 100%/1 | 100%/1 |
| 40 | 25.6855 | 100%/1 | 100%/1 |
| 50 | 25.0753 | 100%/1 | 100%/1 |

as shown in Fig. 4.7(b). The polygons are generated as in Fig. 4.7(c) by the methodology depicted in Fig. 4.1(b). The artificial image in Fig. 4.7(d) on which labeling operation is performed consist of five classes having close mean with additive Gaussian noise of variance $\delta_n^2$. The overlap between classes is high as shown in Fig. 4.7(f) and for better visualization the contrast is stretched in Fig. 4.7(d). The proposed algorithm is applied on the artificial images with different noise levels and the perfect labeling solution in Fig. 4.7(e) is obtained. For each noise level the same test has been performed 10 times and the results are summarized in Table 4.3. Results indicate the robustness of proposed method against different peak signal-to-noise ratio (PSNR) levels. In this case the proposed method with GMM yields the perfect results, since classes are ideally Gaussian distributed.

## 4.4 Summary of Results

Real word images such as images of nature and the SAR images, presented in next chapters, are difficult scenes having variability of class statics and other factors (Section. 1.1), which deviate classes from normal distribution. Feature statistics solely deficient in characterizing the method (Section 4.2). The spatial interactions need to be taken into account as in the proposed method which robustly performs both in ideal and challenging scenarios.



(a) (b) (c)

(d) (e) (f)

Fig. 4.7: (a) Reference image with five classes. (b) Unsupervised classification output with $n_r = 64$. Each region is the same-size. (c) Polygon structure with $n_p = 23$ polygons. (d) Artificial image of five classes with mean 20,40,60,80,100 and Gaussian noise with variance $\delta_n^2 = 50$ is added to mean of each class. (e) Labeling output with proposed method $accuracy = 100\%$ $kappa = 1$ averaged over 10 runs. (f) Distribution of overlapped classes.

# Chapter 5

# Sea-ice Classification

In this chapter the evaluation of the bootstrap approach is performed with operational SAR sea-ice images (Section 5.1.2). Compared to Chapter 4 this is a more challenging application of the bootstrap approach (Section 5.2). The unsupervised classification and polygon boundaries are not ideal, however the expert evaluator trusts the unsupervised classification and provides the labels for the regions. A novel performance evaluation framework has been developed to validate the unsupervised classification and labeling of SAR sea-ice images (Section 5.1). 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 comparison (Section 5.1.2). Comparison demonstrates (Section 5.2.3) success of the bootstrap approach which is the first known successful end-to-end process for operational SAR sea-ice image classification.

## 5.1 Performance Evaluation

A consistent challenge in 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. 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 thesis is a performance evaluation framework that has been implemented to guide an expert to generate a reference image. The evaluator (T. Zagon) has used this framework to generate reference images used for testing in this thesis.

### 5.1.1 Study area

The images were selected deliberately by CIS to represent challenging examples for automatic classification. Table 5.1 provides summary information for each image and the geographical locations of images in the Arctic map as shown in Fig. 5.3. All the images are captured during freeze up which is, especially, challenging since the sea-ice properties are in transition.

### 5.1.2 Performance evaluation framework

Fig. 5.1 shows the flow of the algorithm framework. The framework is built within MAGIC [3]. 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 unsupervised classification 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 undergoes new unsupervised classification otherwise the existing unsupervised classification result is used. The evaluator has to decide if the unsupervised classification of the polygon is successful or if egg code parameters must be modified. If the unsupervised classification fails, the evaluator discards the polygon from the reference image and enters an explanation for the failure. Eliminating any unsupervised

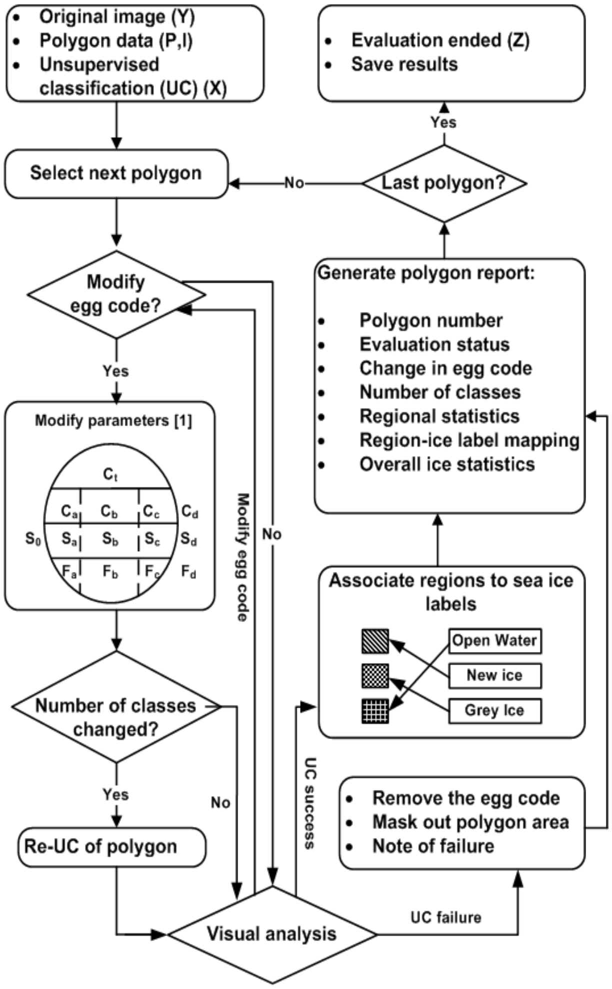


Fig. 5.1: This performance evaluation framework is integrated into MAGIC [3] and is designed for a person to generate full scene classified pixel-level data. The user can modify egg code parameters, reclassify a polygon with new parameters and discard the polygon if unsupervised classification fails. For successfully classified polygons, the user manually assigns a sea-ice label to each region to create the reference database.

classification issues allows direct focus on evaluating the sea-ice labeling performance. The unsupervised classification may fail in some exceptional scenarios where the discrimination depends on characteristics other than $\sigma^o$. Apart from feature failures, the evaluator can discard a polygon if a portion of 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 unsupervised classification usually succeeds and then the evaluator assigns a class label to each region which acts as a reference. A 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.

### 5.1.3 Results of performance evaluation framework

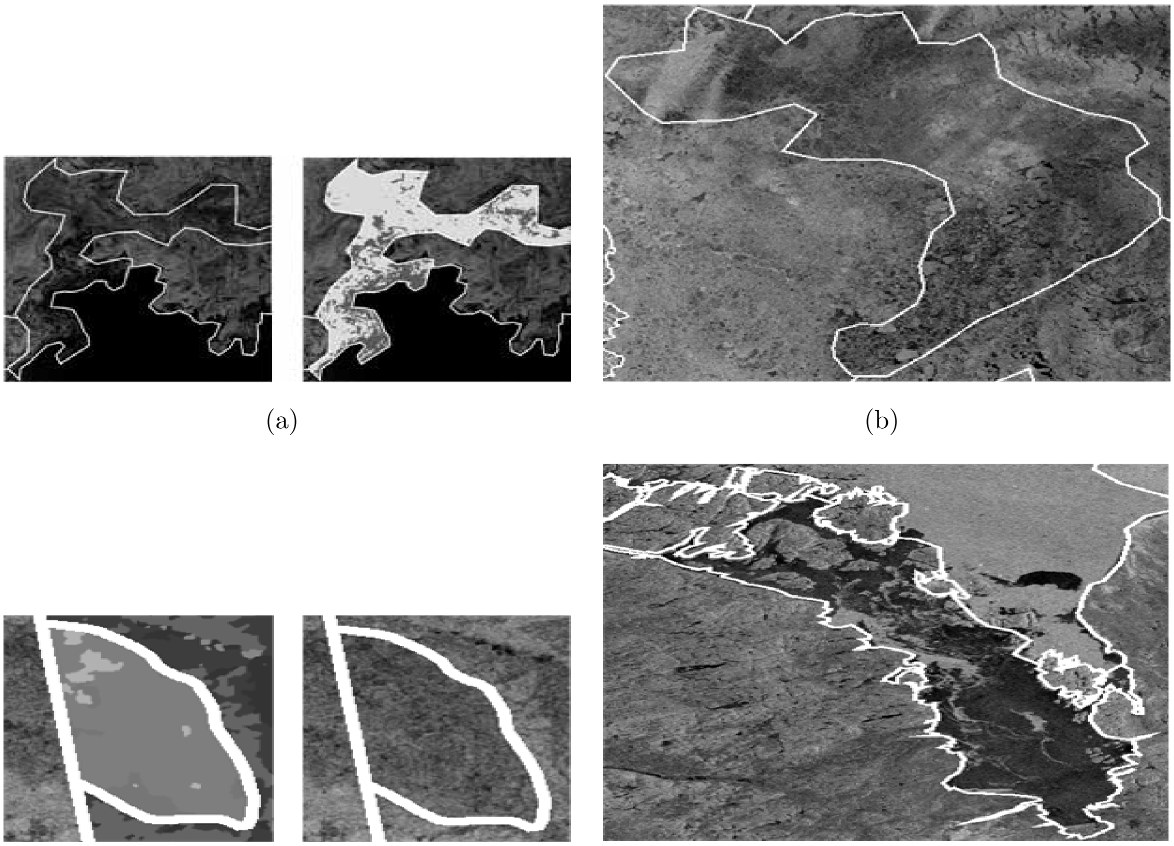
Table 5.1 shows the results of the performance evaluation for each image including how many polygon egg codes were modified and how many polygons were reclassified, discarded, and labeled. The evaluator had to make many changes to the CIS-provided source data to produce accurate reference data.

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 reclassified.

Some polygons were discarded. For Image 1, one polygon (Fig. 5.2(a)), according to the evaluator, requires a fine grey and new ice unsupervised classification. The other Image 1 polygon (Fig. 5.2(b)) is a complicated scene that the evaluator was not able to visually interpret. For Image 2, the three erroneous polygons each have overly smoothed unsupervised classifications and an example is shown in Fig. 5.2(c). Note that all polygons are classified 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. 5.2(d)). The evaluator excluded the polygons to remove any ambiguity.

## 5.2 Classification of Sea-Ice Images: Bootstrap Approach

Thousands of SAR sea-ice images are systematically processed every year in support of operational activities such as ship navigation and environmental monitoring. Currently,



(a) (b)

(c) (d)

Fig. 5.2: Discarded polygons in reference images. (a) Image 1 polygon. According to evaluator, requires a fine grey and new ice unsupervised classification.(b) Another Image 1 polygon. A complicated scene that the evaluator was not able to visually interpret. (c) One of the three Image3 polygons, which had overly smoothed unsupervised classifications. (d) One of the Image 3 five polygons, which were discarded due to errors in the provided land maps where land boundaries were not properly provided and islands were included.

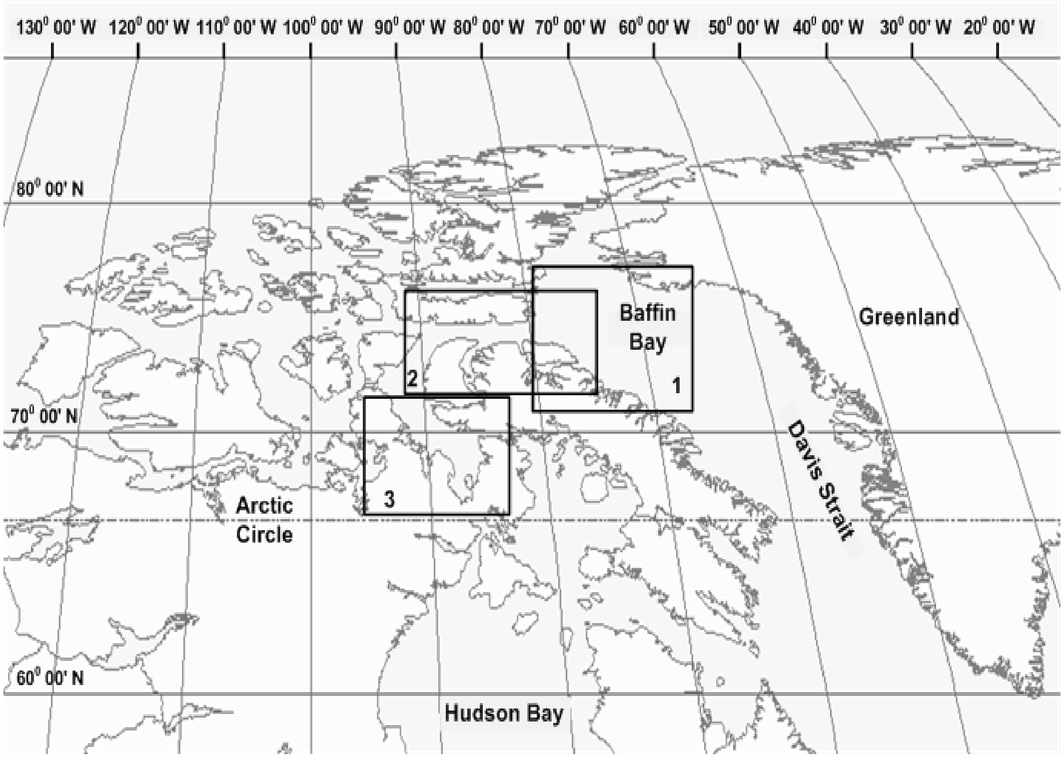


Fig. 5.3: Location of three reference images on map (generated by author using GIS tools).

Table 5.1: Performance evaluation of SAR sea-ice images.

| | Image 1 | Image 2 | Image 3 |
|---|---|---|---|
| Image Area | Baffin Bay | Baffin Bay | Gulf of Boothia |
| Date Acquired | Oct. 30/05 | Oct. 18/05 | Oct. 06/04 |
| Modified egg code | 12 | 13 | 13 |
| Reclassified polygons | 6 | 7 | 7 |
| Discarded polygons | 2 | 3 | 5 |
| Labeled polygons | 12 | 13 | 13 |
| Total polygons | 14 | 16 | 18 |

using a standardized approach, trained ice analysts manually segment full SAR scenes into smaller polygons to record ice types and concentrations. Using this data, classification with bootstrap approach has been achieved by initial unsupervised classification of each polygon followed by automatic sea-ice labeling.

### 5.2.1 Evaluation and polygon data

Reference images produced in Section 5.1 have been used as a basis for performance analysis. The performance metrics in Section 4.1.3 are used for comparison. Table 5.2 summarizes the final polygon data with corresponding ice types used for classification with bootstrap approach.

### 5.2.2 Sea-ice polygon unsupervised classification

IRGS is the algorithm used for unsupervised classification. An example of an IRGS unsupervised classification is shown in Fig. 5.4(b). This polygon is downsized 25 times to fit so details are lost in the rescaling. This shows a typical challenging unsupervised classification of grey, grey-white and new ice types which is not feasible to classify manually given the difficult and time consuming nature of manual unsupervised classification. IRGS is able to successfully classify this polygon and shows similar performance for polygons summarized in Table 5.2.

### 5.2.3 Sea-ice labeling

Table 5.3 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. 5.5 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 (5 Kx5 K) to be shown at full resolution and therefore the WMO color coded results are presented to visualize the outputs.

Table 5.2: Final polygon data for reference database. Using Table 2.1, 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 | $\{5,4,1\}$ | $\{8.,4,W\}$ |

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. 5.5(c)) circled and marked with "X". Intuitively, this polygon could have been perfectly labeled 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 inter-and intra-class feature variability, and with such data spatial

Table 5.3: 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 |

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. 5.5(f)) has seven ice labels found in only 13 polygons. One polygon inside a white circle on the far left of the image in Fig. 5.5(e) 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 labeled 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 between polygons) then the unsupervised classification has a stronger accuracy and the labeling process produces a more accurate pixel-level classified map.

### 5.2.4 Role of prior model

To demonstrate the role of the prior term, a comparison between using both feature and prior terms (Fig. 5.5(a), duplicated for side-by-side viewing in Fig. 5.6(a)) and just the feature term (Fig. 5.6(b)) is performed. Here, the feature model alone is not sufficient to accurately label the regions. By removing the spatial prior edge model, only the proposed method with GMM remains. The labeling result using GMM alone (Fig. 5.6(b)) with

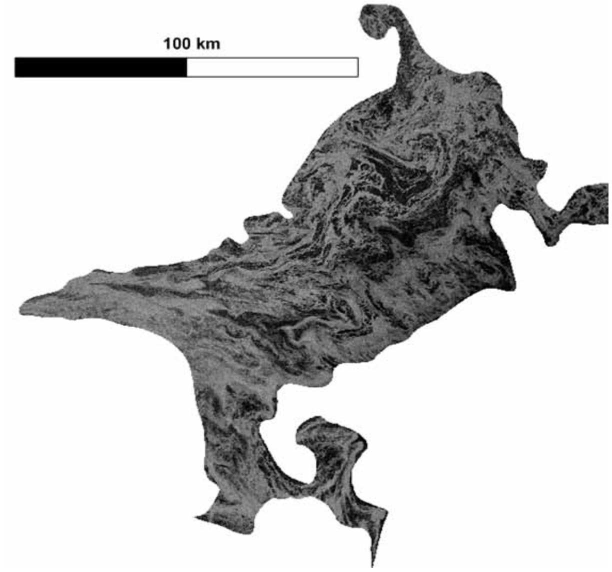
Table 5.4: Computational time of the proposed labeling algorithm.

| | Image 1 | Image 2 | Image 3 |
|---|---|---|---|
| Iterations$\theta$ | 100 | 100 | 100 |
| Regions $n_r$ | 29 | 36 | 32 |
| Polygons $n_p$ | 12 | 13 | 13 |
| Size $MxN$ | 5028x5387 | 4994x5417 | 3818x4688 |
| Time (seconds) $s$ | 31s | 26s | 77s |

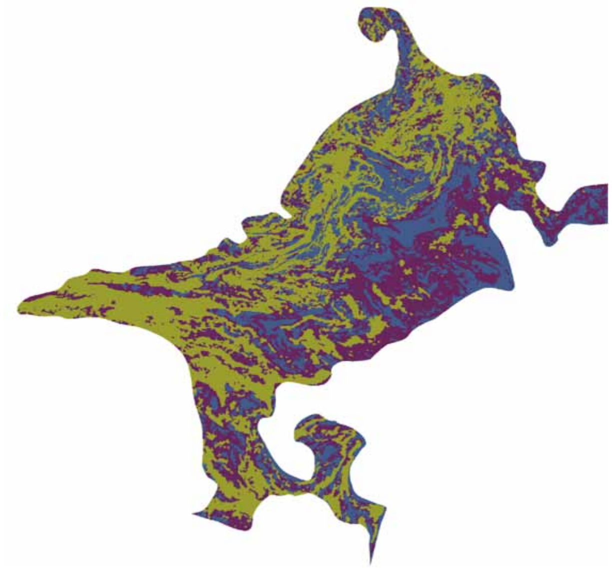
accuracy 33.10% and kappa 0.1059 shows more then half of the regions have been misclassified compared to using both feature and spatial term (Fig. 5.6(a)) with accuracy/kappa 100%/1. This clearly indicates that the spatial interaction of polygons is essential in the overall model to generate accurate labeling.

### 5.2.5 Computational time

The overall complexity of just the labeling algorithm is $O(\theta n_r n_p + MN)$. Table 5.4 summarizes the computational time required to achieve labeling results for every test image after the unsupervised classification. The unsupervised classification 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 algorithm's use in an operational environment.

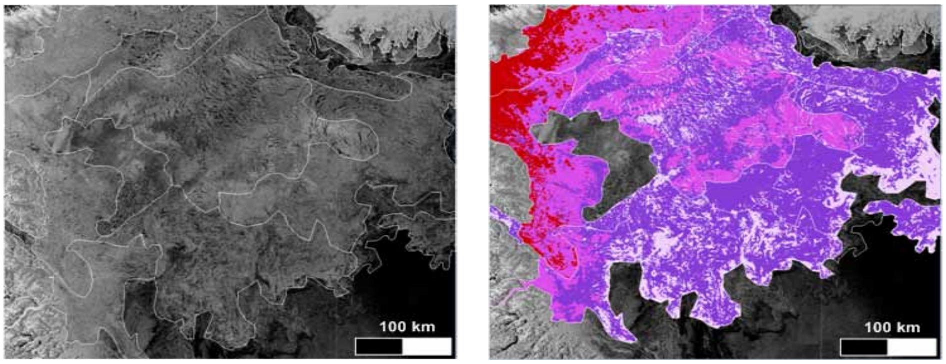


(a)

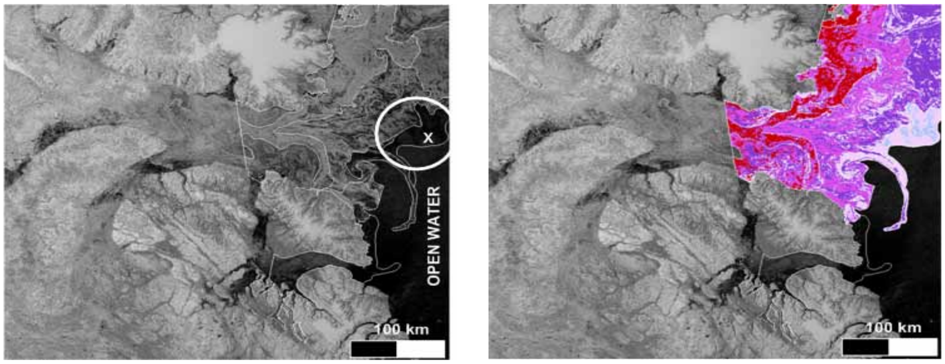


(b)

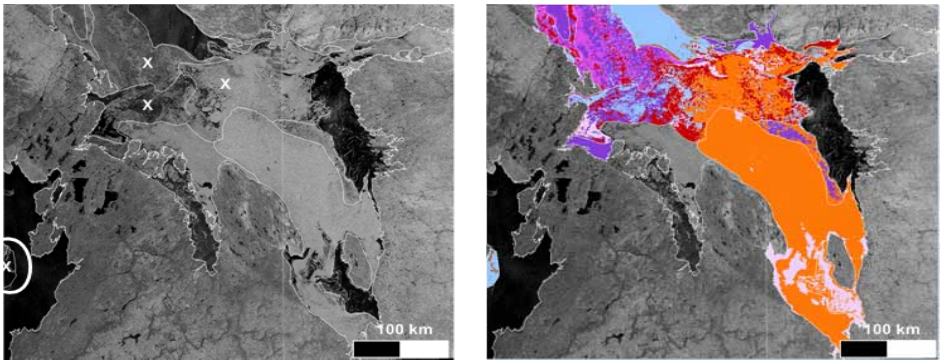
Fig. 5.4: Example of unsupervised classification using IRGS [2] (arbitrary colors used). (a) This Image 2 polygon is approximately 2 Kx2 K pixels and has been downsized 25 times to fit page which removes scene details. The polygon's egg code indicates presence of grey-white, grey and new ice. (b) The polygon in (a) successfully classified to three classes using IRGS [2]. The unsupervised classification performance of such complex imagery demonstrates the capability of IRGS as the unsupervised classification technique.



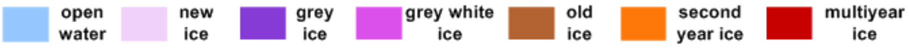
(a) Image 1 (b) Classified Image 1



(c) Image 2 (d) Classified Image 2

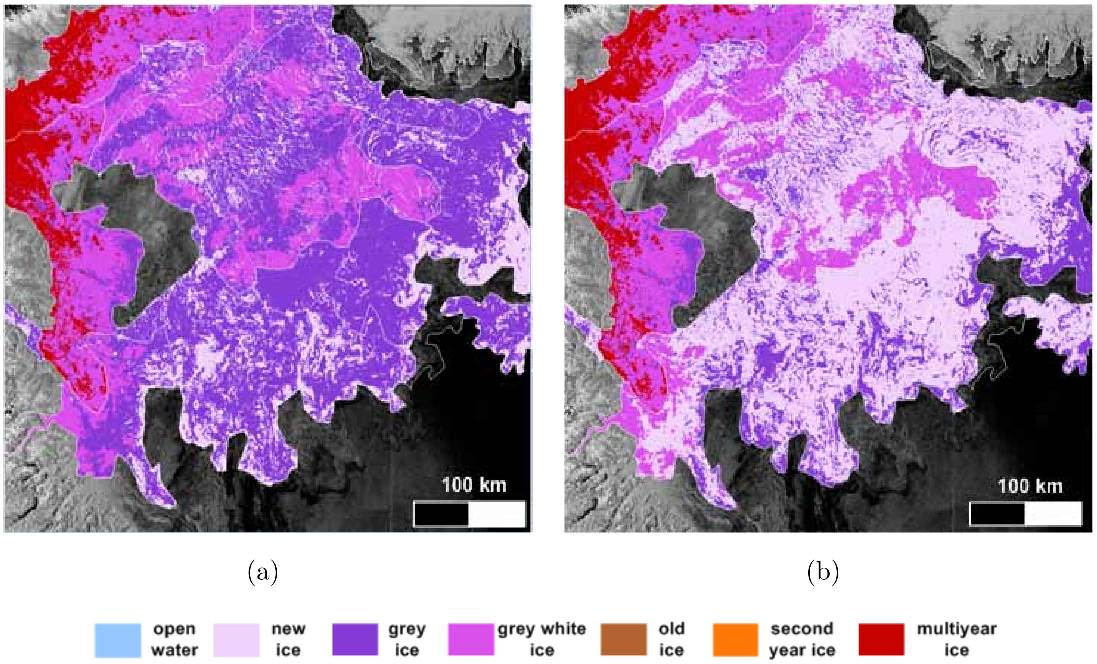


(e) Image 3 (f) Classified Image 3



(g) WMO [1] color code

Fig. 5.5: Classification of SAR sea-ice images with proposed technique. Polygon boundaries are overlaid on the image as white contours. Images are too large (5 Kx5 K) to show details but this unsupervised classification and labeling success can be observed.



(a)

(b)

(c) WMO [1] color code

Fig. 5.6: The effect of spatial term on classification performance of Image 1. (a) The labeled output using the proposed method 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.

# Chapter 6

# Lake Ice Classification

Many northern lakes are monitored for ice conditions for aiding transportation, predicting weather, modeling climate and studying lake dynamics. The manual estimation of ice fraction becomes a tedious process if more frequent and pixel level results are required. The potential of using automated algorithms for SAR lake ice image interpretation is explored using unsupervised classification techniques (Section 6.5.1). The methodology is presented for obtaining reference images used for comparison (Section 6.4). The results of a study have been demonstrated using a series of operational SAR images of Great Bear Lake (GBL) and Great Slave Lake (GSL) (Section 6.5.2). The classification of SAR lake ice images with conventional and bootstrap approaches is performed next. Compared to standard classification approaches, high accuracy has been obtained using the bootstrap approach for difficult SAR lake ice images (Section 6.5.3). The framework used for GBL/GSL can be easily extended to other lakes and can be adapted for operational lake ice monitoring.

## 6.1 Study area

### 6.1.1 Great Bear Lake (GBL)

GBL is located in the Northwest territories (NWT) (66°N 121°W) and is the largest lake lying entirely within Canada and is the eighth largest in the world. The lake is deep,

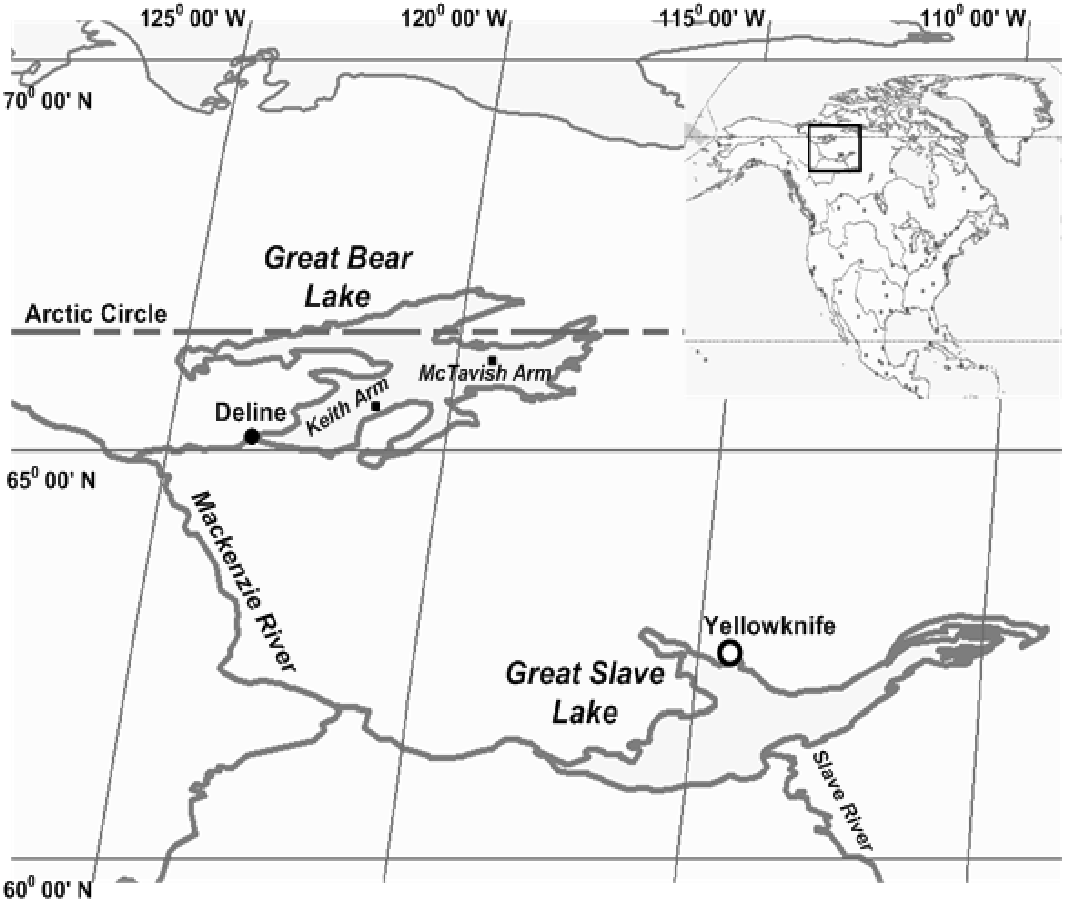


Fig. 6.1: Location of GBL and GSL in the NWT, Canada (generated by author using GIS tools).

reaching a maximum of 446 meters and has a surface area of 31.153 $km^2$ [130]. As seen in Fig. 6.1, portions of the eastern arms of the lake fall inside the Arctic circle and are subject to polar darkness. There are 26 main islands totalling 759.3 $km^2$ [131] (2.5% of the total area) and these must be accounted for in unsupervised classification of fine resolution imagery. The closest meteorological weather station is Deline. GBL is covered with ice from November to July and has one ice road connecting to the community in Deline, NWT.

### 6.1.2 Great Slave Lake (GSL)

GSL (61°N 114°W) is also located in NWT at a slightly lower latitude than Great Bear Lake as shown in Fig. 6.1. GSL covers 27.200 $km^2$ area and is the next largest lake in that region [130]. GSL is also the deepest lake in North America with the maximum recorded depth of 614 m. The ice on the lake stays for eight months (November-June) each year and the ice road connecting Yellowknife to Detah is open in the winter time. Yellowknife is the main weather station nearby the lake.

## 6.2 Ancillary Data

Meteorological station data were obtained from online archives of Environment Canada [132]. Deline and Yellowknife, the major weather stations, have been selected as representative of GBL and GSL temperatures due to their vicinity to the lakes and the completeness of information. The data contain the daily mean temperatures for years 1997-2007. Since the period of RADARSAT images is weekly, the change in ice fraction values has been compared to weekly mean temperatures.

## 6.3 SAR Lake Ice Image Processing

### 6.3.1 Lake ice phenology events

Lake ice phenology events are used as major stages of ice formation, therefore defined here. Ice phenology events occur in sequential order. The melting occurs on the first day open water seen on the lake and is referred to as melt onset (MO). The ice continues to melt during the break up period until the water is clear of ice (WCI). Freeze onset (FO) is the first date ice is seen on the lake followed by ice growth in the freeze up period until the day of complete freeze over (CFO) is reached.

### 6.3.2 SAR imaging of lake ice

Although similar in some aspects SAR lake ice imaging differs from SAR sea and river ice imaging. SAR lake ice imaging can be further categorized as SAR imaging of shallow and deep lakes. The extensive bubble formation and freezing to ground are the main characteristics of shallow lakes; however, the focus of this study is on deep lakes where the ice remains afloat. This section introduces some of the concepts of lake ice formation from past studies (1991-1992) [54] and from available images acquired in 1998. The purpose of this introduction is to comprehensively analyze the unsupervised classification outputs in Section 6.5.1.

Thirteen sites were observed on McTavish Arm of GBL using the ERS-1 SAR data years 1991-1992 [54]. For every given date the average backscatter intensity has been taken of those six locations and depicted as squares in Fig. 6.2. Morris et al. [54] indicates that the ice cover reaches a stable establishment in late December, then the backscatter continually rises after it reaches a maximum in the January when the level holds for the rest of the winter. To complement the graph, a region has been selected from Keith Arm on GBL RADARSAT-1 images of June-December 1998 and its average backscatter intensity for every date is depicted as triangles in Fig. 6.2. The values characterize ice in the late melting season and its formation before the stable cover is established. This is the season when ice fraction estimation matters.

Regions in Fig. 6.3 are depicting the same location selected from Keith Arm of GBL for different dates. Visually and based on dates the relation will be made with backscatter characteristics of SAR lake ice imagery. At final break up period, day 179, the lake is ice free and most of the radar signal is reflected in the opposite direction producing very low return of average -23dB. Freshwater has a high dielectric constant of $\varepsilon_{FW} = 80$ and behaves as an ideal reflector in calm weather. In the middle of November, day 323, the lake surface is mostly slush with reflected signal scattered by particles, and high return is received at the antenna, around -17 dB. Next week, day 330, new clear ice is starting to form which appears as dark [32]. New clear ice is bubble free and very minimal backscatter returned with most of the signal reflected from ice water surface due to the high difference in dielectric constants (bubble free freshwater ice $\varepsilon_I = 3.17$ [133]). At this stage strong winds can break and displace the thin ice producing a rough surface with cracks, ridging and rafting features which have sharp corners [32]. That is why some bright contours can

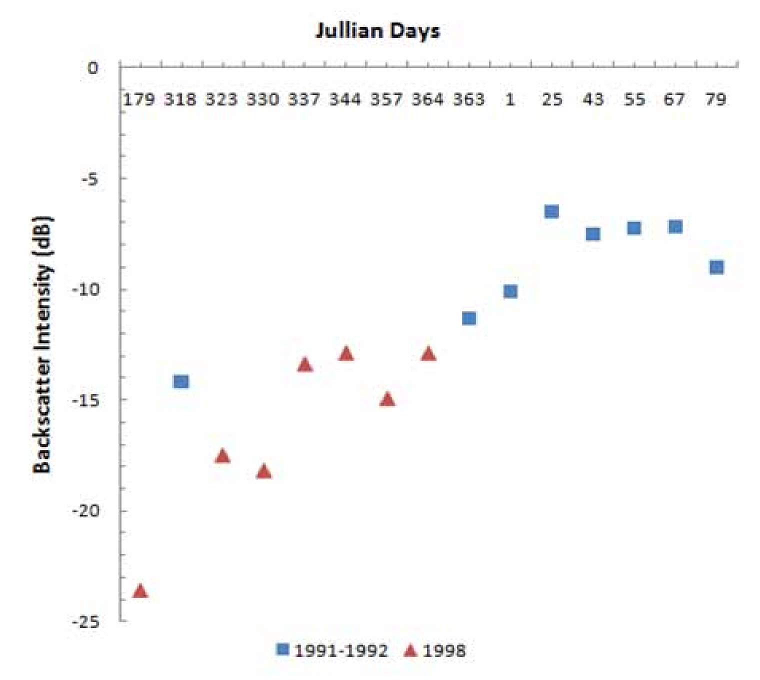


Fig. 6.2: Evolution of lake ice backscatter over the time. The values shown as squares were obtained by averaging the backscatter intensities over ERS-1 image samples from the eastern side (McTavish Arm) of GBL on days during November 1991 to May 1992 [54]. The values shown as triangles were obtained by averaging the backscatter intensities over RADARSAT-1 image samples from western side (Keith Arm) of GBL on dates during June-December 1998.

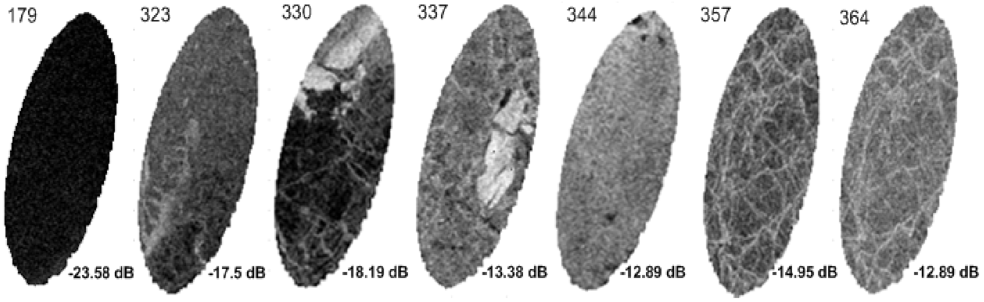


Fig. 6.3: Temporal backscatter characteristics of SAR imaging of the deep lake (GBL 98) from open water to complete freeze over. Presented to visually assess the variability of lake ice features.

be seen on the same image which corresponds to the rough surface areas. A week later, day 337, as ice thickens the air bubble inclusion increases and the signal reflected from the ice freshwater surface bounces from those bubbles and increases backscatter significantly, in this case almost -5dB increase. The effect of bubbles is more profound in some of the shallow lakes [32, 55, 56] due to the presence of gases such as methane near the surface of the lake [134]. Ice continues to thicken and consolidate as seen at day 344 and the backscatter increases, probably, with some ratio related to thickness. Apart from ice thickness/bubbles, ice stratification, snow ice and the conditions at ice/water interface are also important scattering parameters in C-band [29, 55, 109] and might have affected the backscatter for day 357. Large accumulation of the snow apply pressure that cause the ice to crack with some water appearing on the surface [56, 135]. Later, during the next week the water from the cracks might have combined with snow and formed the snow ice that contributed to higher backscatter for day 364.

## 6.4 Performance Evaluation

### 6.4.1 Author generated reference images

The experiments have been conducted by applying the methods discussed in Section 2.4 to seven SAR lake ice images of GBL/GSL. Three of those images were acquired in year 2004 and the rest in 1998. The reference images have been manually produced for all seven images since fully validated field reference for the SAR lake ice images is not available. To obtain pixel level validated images of GBL/GSL, one would have to perform the field sampling of the lake ice on site across hundreds of kilometers during each SAR satellite overpass. Such a validation exercise is logistically impossible, instead the author manually created three pixel level reference images from SAR lake ice images using the basics of lake ice formation (Section 6.3.2).

Three such images have been produced by pixel level detailed analysis of SAR lake ice images. The availability of two cloudless MODIS images for melting season of 2004 has also helped to guide the reference image generation. Section 6.5.1 presents the unsupervised classification performance by comparing the results with these three reference images.

### 6.4.2 Blind test

The external expert and evaluator (Nic Svacina) was also asked to produce the manual classification results with the sequence of images from year 1998. The evaluator had no familiarity and exposure to algorithmic results. The purpose of blind test was to eliminate the possible bias with the previous approach and to better asses the performance of classification techniques. The four reference of SAR lake ice images were obtained by the evaluator who provided pixel-level classified reference images using manual GIS tools. The performance figures obtained with such reference images are analyzed in the second part of Section 6.5.1.

### 6.4.3 Performance metrics

The reference images were used as a basis for evaluating the lake ice unsupervised classification and labeling techniques. The performance metrics in Section 4.1.3 are used for comparison. Furthermore, using the kappa coefficient in Eq. 4.1 and its standard error $\sigma_\kappa$ significance testing for unsupervised classification methods was performed and results presented as a significance matrix. The diagonal values in the matrix depict the significance of testing of individual methods obtained from equation:

$$Sig = \frac{\kappa}{\sigma_\kappa} \tag{6.1}$$

where $Sig$ is normally distributed [136], whereas the off-diagonal values are obtained from pairwise statistics:

$$Sig = \frac{\kappa_1 - \kappa_2}{\sqrt{\sigma_{\kappa_1}^2 + \sigma_{\kappa_2}^2}} \tag{6.2}$$

where standard error of $\kappa$ given as [129]:

$$\sigma_\kappa = \sqrt{\frac{P(A)(1-P(A))}{N(1-P(E))^2}} \tag{6.3}$$

For a 95% confidence level if $Sig$ is greater than the 1.96 threshold the unsupervised classification results are significant and better than a random unsupervised classification. Similarly when compared to the same threshold a pairwise significance test indicates how the two unsupervised classification methods are significantly different.

## 6.5 Classification of Lake Ice Images

### 6.5.1 Unsupervised classification of SAR lake ice images

**Comparison with author generated reference images**

One of the images used for comparison is the SAR image of GBL (Fig. 6.4) acquired during the melting season (July 2, 2004). The backscatter intensity expected from open water in this season is generally very low (Section 6.3.2) making open water clearly distinguishable as seen in Fig. 6.4(a). In the same figure, the brighter lake areas refer to consolidated ice with high backscatter. The challenge during the melting season is the presence of dark grey areas of ice, probably at an advanced stage of melt. These areas in Fig. 6.4(a) are located in northern part of the lake and can be easily misinterpreted as cracks with open water. However, in counterpart, for the MODIS optical image in Fig. 6.4(b) the same areas are clearly seen as being ice covered. This information is reflected in the reference image shown in Fig. 6.4(c). The unsupervised classification results with IRGS given in Fig. 6.4(d) indicate a higher performance compared to other methods since fewer pixels of dark grey ice is merged with the open water region (Table 6.1).

The same experiment was conducted with the SAR imagery of GSL in the same season (June 16, 2004). Similarly, the challenge in Fig. 6.5(a) is the darker grey ice in the eastern arm. The methods such as GMM and K-means, are susceptible to such scenarios since they do not account for spatial locality. Both of the methods merged the dark grey ice pixels as open water and produced lower metric of kappa compared to MRF based techniques C-MLL and IRGS which have near perfect performance in Table 6.1. The unsupervised classification of SAR lake ice images in Fig. 6.6(a) is nontrivial which corresponds to the freeze up period. As discussed in Section 6.3.2, the newly formed ice appears dark in SAR imagery and can also look bright due to fractures and ridges. The dark areas with ice can be easily misinterpreted as open water by clustering techniques, K-means and GMM, which exhibit low performance as seen in Fig. 6.6(c,d) and Table 6.1. The GMM, in addition, misclassified the roughed water on the western side as ice. The essential property required to efficiently segment this image is utilizing the strong edges at the boundary of open water and ice. Such a criteria does not exist in the spatial term of C-MLL which misclassified the newly formed ice with open water as seen in the Fig. 6.6(e). The IRGS has a substantial success compared to other techniques in Table 6.1 and Table 6.2 since it combined most of

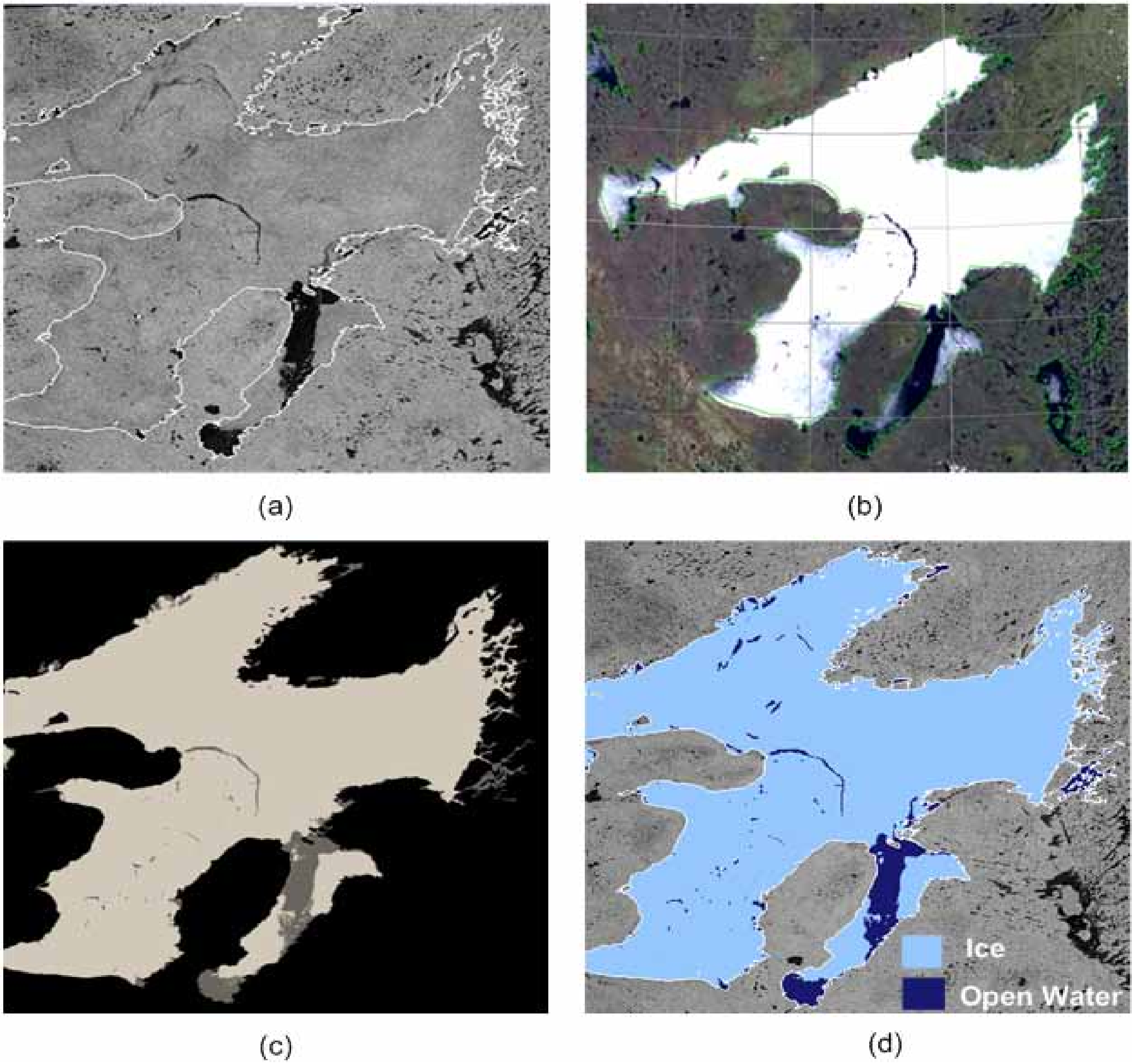


Fig. 6.4: IRGS unsupervised classification result compared to reference obtained using SAR imagery of GBL July 2, 2004. (a) Original SAR image. (b) MODIS image of GBL July 2, 2004. (c) Reference image. (d) Unsupervised classification result with IRGS

Table 6.1: Performance of unsupervised classification algorithms applied to SAR lake ice imagery of GBL/GSL 2004 (accuracy/kappa). Standard error of $\kappa$ ($\sigma_k * 10^{-3}/\kappa$) is given in second table. The example of confusion matrix for GSL Nov, 18 is given in third table. Diagonal values represent the number of pixels correctly classified for each class. Off-diagonal values indicate the number of misclassified pixels (open water misclassified as ice (e.g K-means 2857) and ice misclassified as open water (e.g K-means 614052)). Significance test of the statistics is provided in Table 6.2

| Methods | GSL Jun 16 | GBL Jul 2 | GSL Nov 18 |
|---|---|---|---|
| K-means | 97.5%/0.92 | 98.8%/0.90 | 71.2%/0.44 |
| GMM | 98.3%/0.94 | 98.8%/0.91 | 72.0%/0.43 |
| C-MLL | 98.4%/0.95 | 99.2%/0.94 | 81.8%/0.64 |
| IRGS | 98.5%/0.95 | 99.4%/0.95 | 89.8%/0.80 |

Standard error of $\kappa$ ($\sigma_k * 10^{-3}/\kappa$).

| Methods | GSL Jun 16 | GBL Jul 2 | GSL Nov 18 |
|---|---|---|---|
| K-means | 0.30/0.92 | 0.52/0.90 | 0.61/0.44 |
| GMM | 0.27/0.94 | 0.52/0.91 | 0.62/0.43 |
| C-MLL | 0.27/0.95 | 0.42/0.94 | 0.50/0.64 |
| IRGS | 0.26/0.95 | 0.38/0.95 | 0.41/0.80 |

The example of confusion matrix for GSL Nov 18.

| K-means | | open water | ice |
|---|---|---|---|
| | open water | 1028800 | 614052 |
| | ice | 2857 | 499222 |

| GMM | | open water | ice |
|---|---|---|---|
| | open water | 511149 | 80505 |
| | ice | 520508 | 1032769 |

| C-MLL | | open water | ice |
|---|---|---|---|
| | open water | 1025683 | 384205 |
| | ice | 5974 | 729069 |

| IRGS | | open water | ice |
|---|---|---|---|
| | open water | 1024139 | 211731 |
| | ice | 7518 | 901543 |

Table 6.2: Significance test of the statistics obtained in Table 6.1. The significance values were computed for each method and image given in Table 6.1. The values in diagonals are obtained using Eq. 6.1 and off diagonal values are obtained using Eq. 6.2. Since matrices are symmetric, only the lower triangle is shown. All the figures are above 1.96 for confidence percentage of 95%, therefore all the statistics in Table 6.1 are significant. The example of confusion matrix for GSL Nov 18 is given in Table 6.1 and the standard error of $\kappa$ and $\kappa$ used to compute the significance values are given in Table 6.1

| GSL Jun 16 | K-means | GMM | C-MLL | IRGS |
|---|---|---|---|---|
| K-means | 3090 | | | |
| GMM | 51.7 | 3556 | | |
| C-MLL | 58.4 | 6.6 | 3650 | |
| IRGS | 62.0 | 10.2 | 3.6 | 3701 |
| GBL Jul 2 | K-means | GMM | C-MLL | IRGS |
| K-means | 1725 | | | |
| GMM | 3.8 | 1751 | | |
| C-MLL | 50.6 | 46.8 | 2211 | |
| IRGS | 71.1 | 67.3 | 21.0 | 2479 |
| GSL Nov 18 | K-means | GMM | C-MLL | IRGS |
| K-means | 719.9 | | | |
| GMM | 7.3 | 689.0 | | |
| C-MLL | 255.8 | 259.2 | 1230 | |
| IRGS | 492.4 | 490.9 | 235.5 | 1937 |

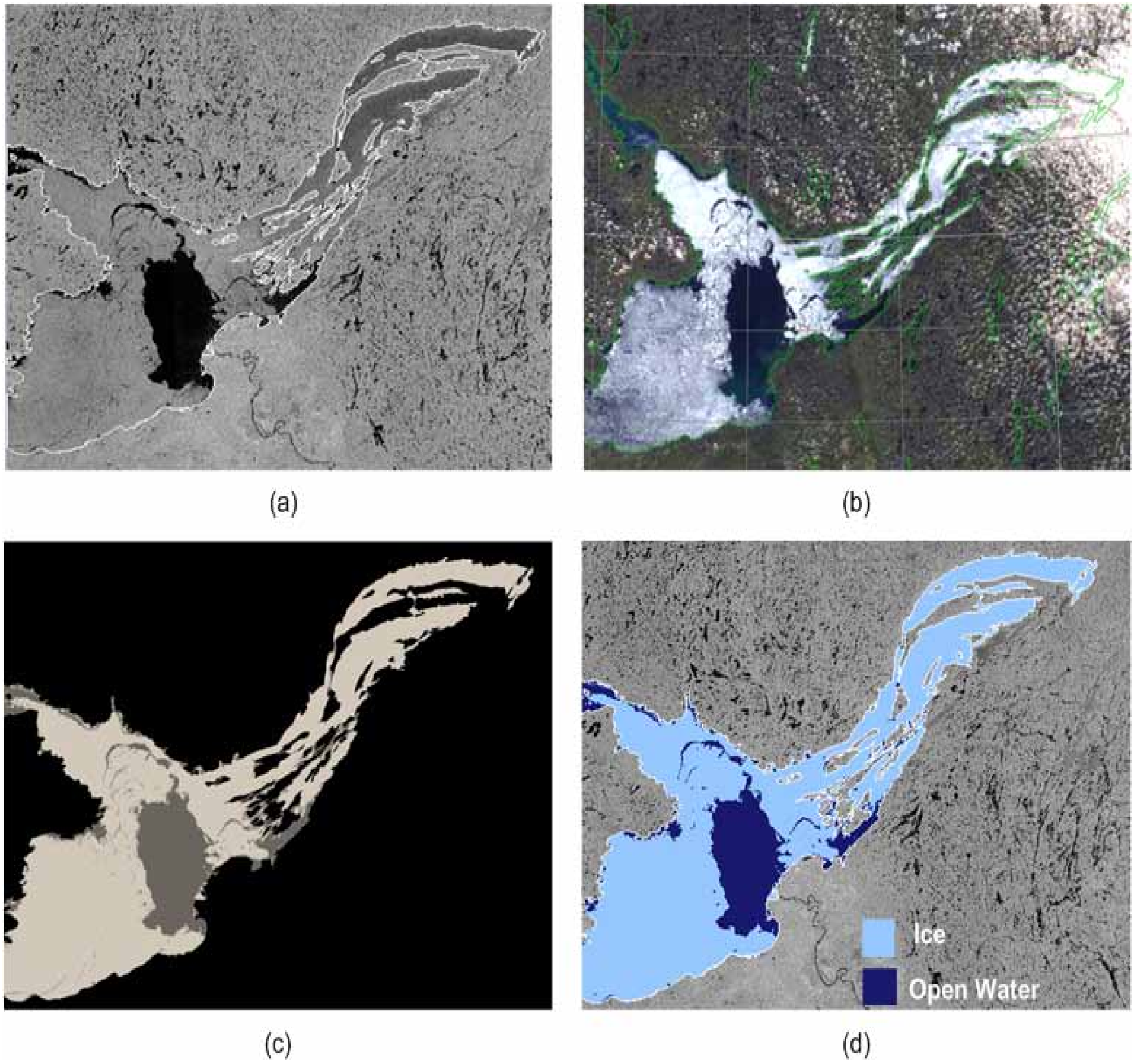


Fig. 6.5: IRGS unsupervised classification result compared to reference obtained using SAR imagery of GSL June 16, 2004. (a) Original SAR image. (b) MODIS image. (c) Reference Image. (c) Unsupervised classification result with IRGS.

the ice in a single region. The only misclassification is seen in the upper part of the center of the lake where the boundaries between the open water and ice are vague.

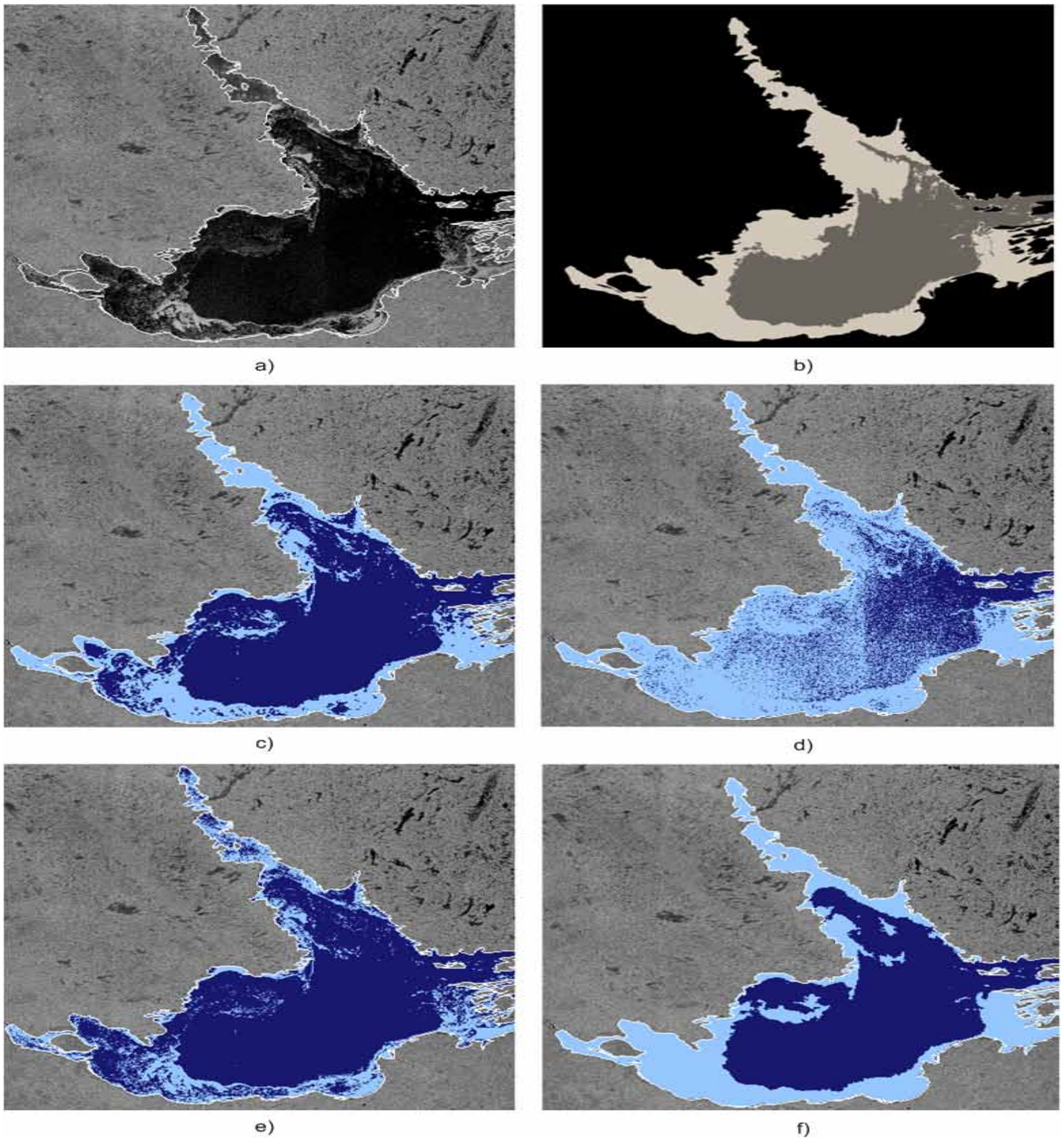


Fig. 6.6: Unsupervised classification techniques compared (accuracy/kappa) to reference obtained using SAR imagery of GSL November 18, 2004. (a) Original SAR image. (b) Reference image. (c) K-means 71.2%/0.44. (d) GMM 72.0%/0.43. (e) C-MLL 81.8%/0.64. (f) IRGS 89.8%/0.80.

### Comparison with reference images generated by blind test

To further assess the performance of SAR lake ice classification, the results in Table 6.3 have been obtained using the reference image provided by the external expert. The results indicate the same high performance of IRGS unsupervised classification for images during the melting season. Compared to other techniques IRGS has a higher kappa for images of the freeze-up period. However, for the same images the kappa values for all unsupervised classification methods are lower because of higher error rate. Nevertheless, the significance values in Table 6.4 indicate that all the unsupervised classification results are significant and significantly differ from each other.

Table 6.3: Blind test evaluation of unsupervised classification algorithms applied to SAR lake ice imagery of GBL/GSL 1998 (accuracy/kappa). Standard error of $\kappa$ ($\sigma_k * 10^{-3}/\kappa$) is given in second table.

| Methods | GBL May 26 | GBL Nov 19 | GSL May 27 | GSL Dec 17 |
|---|---|---|---|---|
| K-means | 94.9%/0.52 | 60.0%/0.08 | 85.3%/0.69 | 57.3%/0.15 |
| GMM | 91.5%/0.40 | 85.3%/0.27 | 87.1%/0.72 | 57.4%/0.15 |
| C-MLL | 97.8%/0.73 | 67.5%/0.15 | 88.0%/0.74 | 57.6%/0.16 |
| IRGS | 97.9%/0.74 | 72.4%/0.28 | 89.3%/0.77 | 64.6%/0.29 |

Standard error of $\kappa$ ($\sigma_k * 10^{-3}/\kappa$).

| Methods | GBL May 26 | GBL Nov 19 | GSL May 27 | GSL Dec 17 |
|---|---|---|---|---|
| K-means | 1.2/0.52 | 0.6/0.08 | 0.5/0.69 | 0.6/0.15 |
| GMM | 1.1/0.40 | 1.0/0.27 | 0.4/0.72 | 0.6/0.15 |
| C-MLL | 1.0/0.73 | 0.7/0.15 | 0.4/0.74 | 0.6/0.16 |
| IRGS | 1.0/0.74 | 0.7/0.28 | 0.4/0.77 | 0.6/0.29 |

Table 6.4: Significance test of the statistics obtained in Table 6.3. The significance values were computed for each method and image given in Table 6.3. The values in diagonals are obtained using the Eq. 6.1 and off diagonal values are obtained using the Eq. 6.2. Since matrices are symmetric only the lower triangle is shown. All the figures are above 1.96 for confidence percentage of 95%, therefore all the statistics in Table 6.3 are significant. The standard error of $\kappa$ and $\kappa$ used to compute the significance values are given in Table 6.3

| GBL May 26 | K-means | GMM | C-MLL | IRGS |
|---|---|---|---|---|
| K-means | 437.7 | | | |
| GMM | 74.5 | 350.0 | | |
| C-MLL | 138.3 | 221.9 | 721.0 | |
| IRGS | 144.3 | 228.4 | 6.0 | 735.7 |
| GBL Nov 19 | K-means | GMM | C-MLL | IRGS |
| K-means | 121.8 | | | |
| GMM | 161.9 | 270.7 | | |
| C-MLL | 79.1 | 96.5 | 220.2 | |
| IRGS | 216.6 | 6.6 | 130.3 | 418.7 |
| GSL May 27 | K-means | GMM | C-MLL | IRGS |
| K-means | 1530 | | | |
| GMM | 51.5 | 165.3 | | |
| C-MLL | 88.1 | 36.2 | 177.4 | |
| IRGS | 130.1 | 78.0 | 42.0 | 1903 |
| GSL Dec 17 | K-means | GMM | C-MLL | IRGS |
| K-means | 246.7 | | | |
| GMM | 2.3 | 249.9 | | |
| C-MLL | 7.9 | 5.6 | 258.4 | |
| IRGS | 171.6 | 169.3 | 163.7 | 500.9 |

### 6.5.2 Manual lake ice labeling

To estimate ice fractions from SAR lake ice images the next stage is to assign proper labels to regions produced from unsupervised classification. In this section the unsupervised approach described in Section 2.6.2 is followed where the regions are manually labeled upon unsupervised classification. The unsupervised approach for SAR lake ice imagery is depicted in Fig. 6.7 and summarized in items below:

I The SAR, shape data discussed in Section 2.2 and the number of classes are the inputs to the system.

II Per requirements discussed in Section 2.2, block averaging is performed for non CIS data.

III The SAR data is automatically converted to $\sigma^\circ$ and ground control points (GCP) are extracted from trailer files to project the boundary of the lake stored as geocoded vector data.

IV In this step, the surface of lakes GBL/GSL is automatically classified using IRGS. Based on image analysis and the number of provided classes only those dates are classified which do not correspond to one of the lake ice phenology events: CFO, WCI, MO, FO (Section 6.3.1). Excluded images are directly passed to the next stage of processing and the others are classified based on number of classes. Ideally, the number of classes has to be two to locate ice and open water, but the image might also contain different ice types requiring it to be greater then two. It is recommended to provide IRGS with maximum number of classes matching the image.

V Ice/no ice labeling of regions is the step manually performed in the unsupervised approach. However, this step can be automated with a bootstrap approach. For instance, the image in Fig. 6.8(a) has been classified to two classes and the open water has been merged with ice in Fig. 6.8(b), whereas, if classified to three the open water would have appeared in a separate region as shown in Fig. 6.8(c). The two other ice regions can be merged manually Fig. 6.8(d). The final result is the ice/no ice labeling with corresponding fraction values.

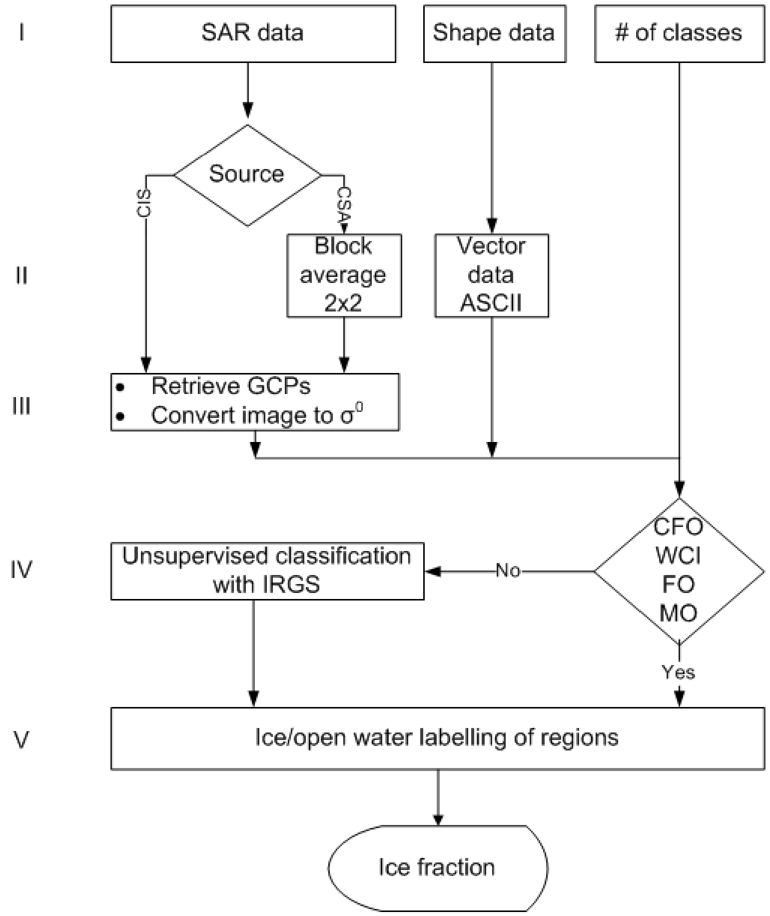
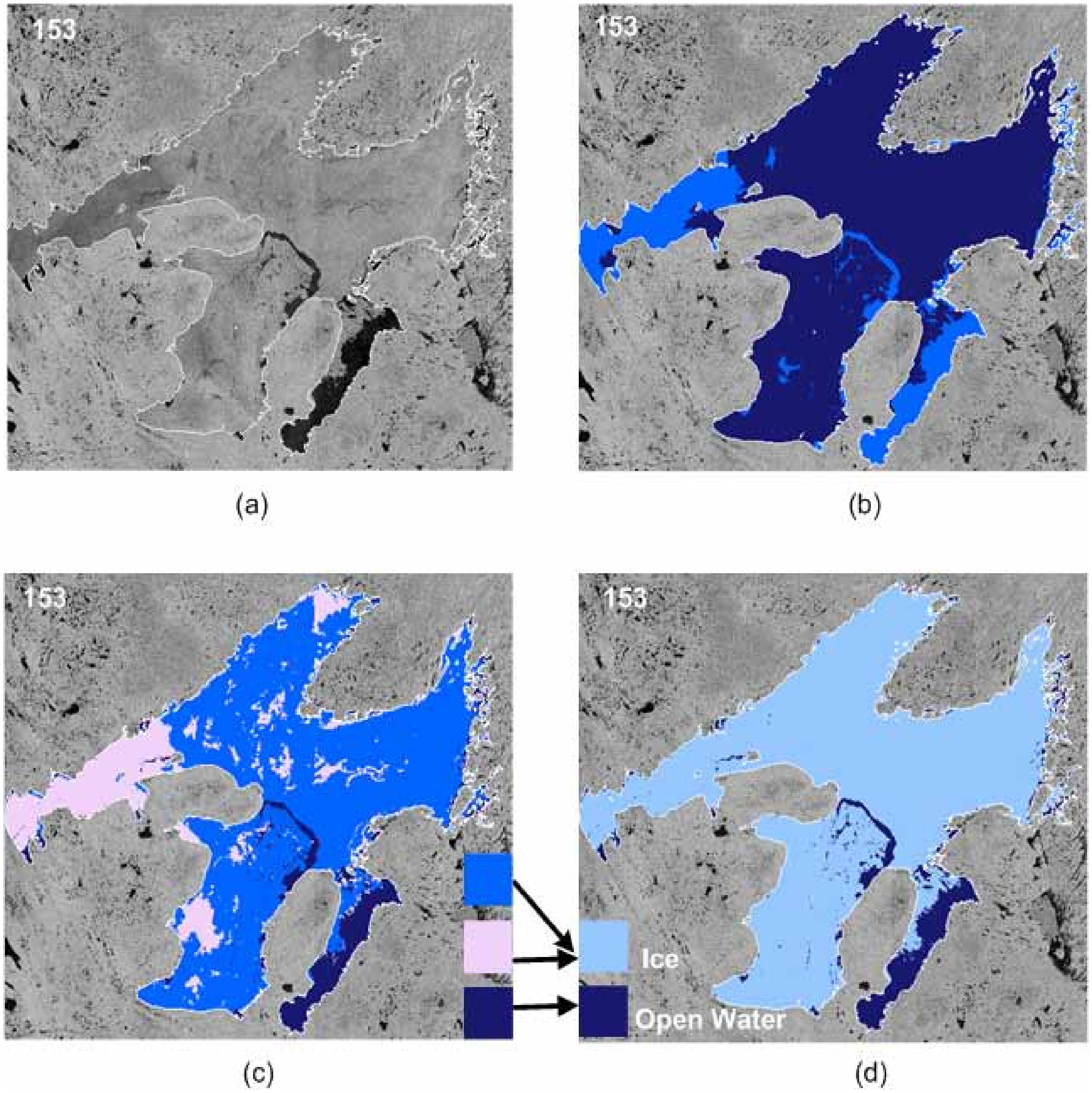


Fig. 6.7: Flowchart of ice fraction estimation with unsupervised approach.

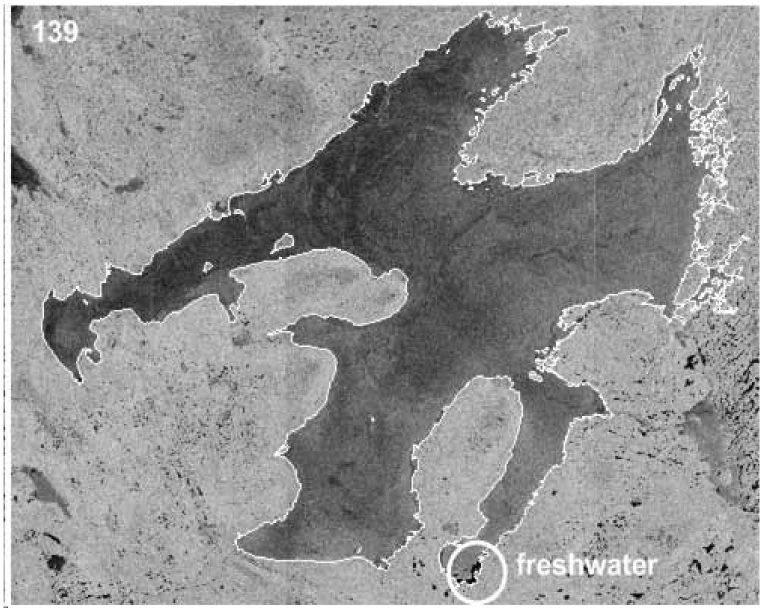
The images for which the number of classes is 1 correspond to CFO, WCI days. They are labeled and assigned ice fractions accordingly, as 10 or 0. For the MO, FO days the presence of little open water/ice compared to the whole lake is insignificant for the unsupervised classification algorithms. Either the smaller region of interest has to be selected for unsupervised classification or the thresholding needs to be done to roughly separate open water from ice. Examples are shown in Fig. 6.9(a,b) for the days 139 and 316 of year 1998 where the little areas of open water and ice are circled in white.



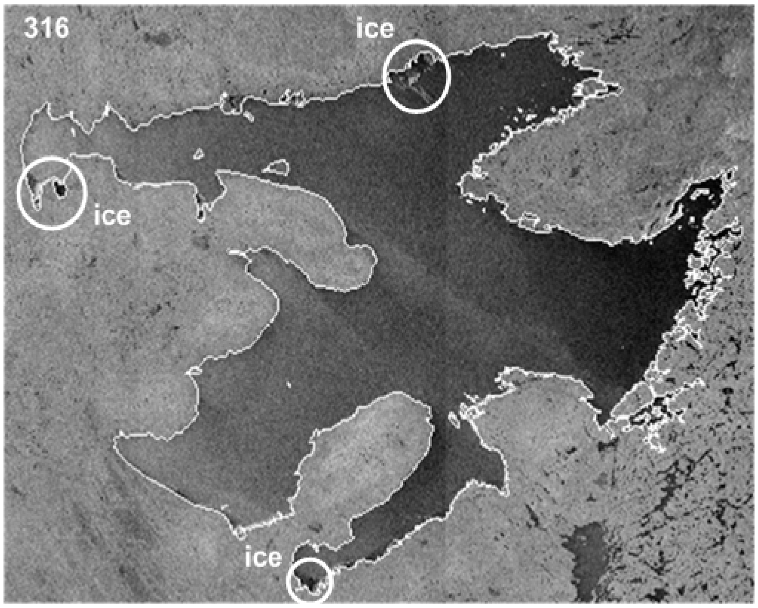
(a) (b)

(c) (d)

Fig. 6.8: Unsupervised classification and subsequent labeling of regions as ice or open water, manually. (a) SAR image June 2, 1998. (b) Two class unsupervised classification where portion of ice has merged with open water. (c) Unsupervised classification with three classes to isolate open water as one class. (d) Two regions are manually merged into same ice class.



(a)



(b)

Fig. 6.9: Start of ice break up/freeze up (MO/FO) period. (a) Very small area in the white circle needs to be classified and identified as open water. (b) Very small areas in the white circles need to be classified and identified as ice.

### Unsupervised approach and CIS Ice Fraction Estimate Comparison

### Great Bear Lake

The dates of temporal images in Fig. 6.10 cover a cycle of lake ice phenology events. All the SAR images in Fig. 6.10 were classified with the methodology described in Section 6.5.2. The set of classified images in Fig. 6.11 shows both the spatial evolution of lake ice as well as its ice fraction values. Looking at the unsupervised classification results in Fig. 6.11 the ice melting can be easily tracked starting from the bottom left arm of the lake and continuing towards the west. It is evident that unsupervised classification results help understanding the lake ice dynamics both spatially and quantitatively in terms of ice fractions.

The extensive results summarized in Table 6.5 are unique in comparing ice fraction estimates from unsupervised approach and CIS. The testing captures eleven years of SAR image acquisitions. Some weeks in Table 6.5 are lacking SAR data or has been omitted due to the partial presence of the lake in the SAR image. The proportions of the lakes are included in the last column of Table 6.5. As seen, the images with less than 0.4 lake area have been omitted. In addition, the images have been avoided which have a very high incidence effect where ice and open water separation is visually ambiguous. The rest of the images have been processed and the results with unsupervised IRGS approach and CIS are reported. The weekly mean temperatures are also included in a column of Table 6.5 to identify the reason of ice fraction change. Generally, lake ice is gradually melting or freezing as a consequence of weekly accumulation of temperature changes, therefore, weekly mean temperature is used to correlate with ice fraction changes.

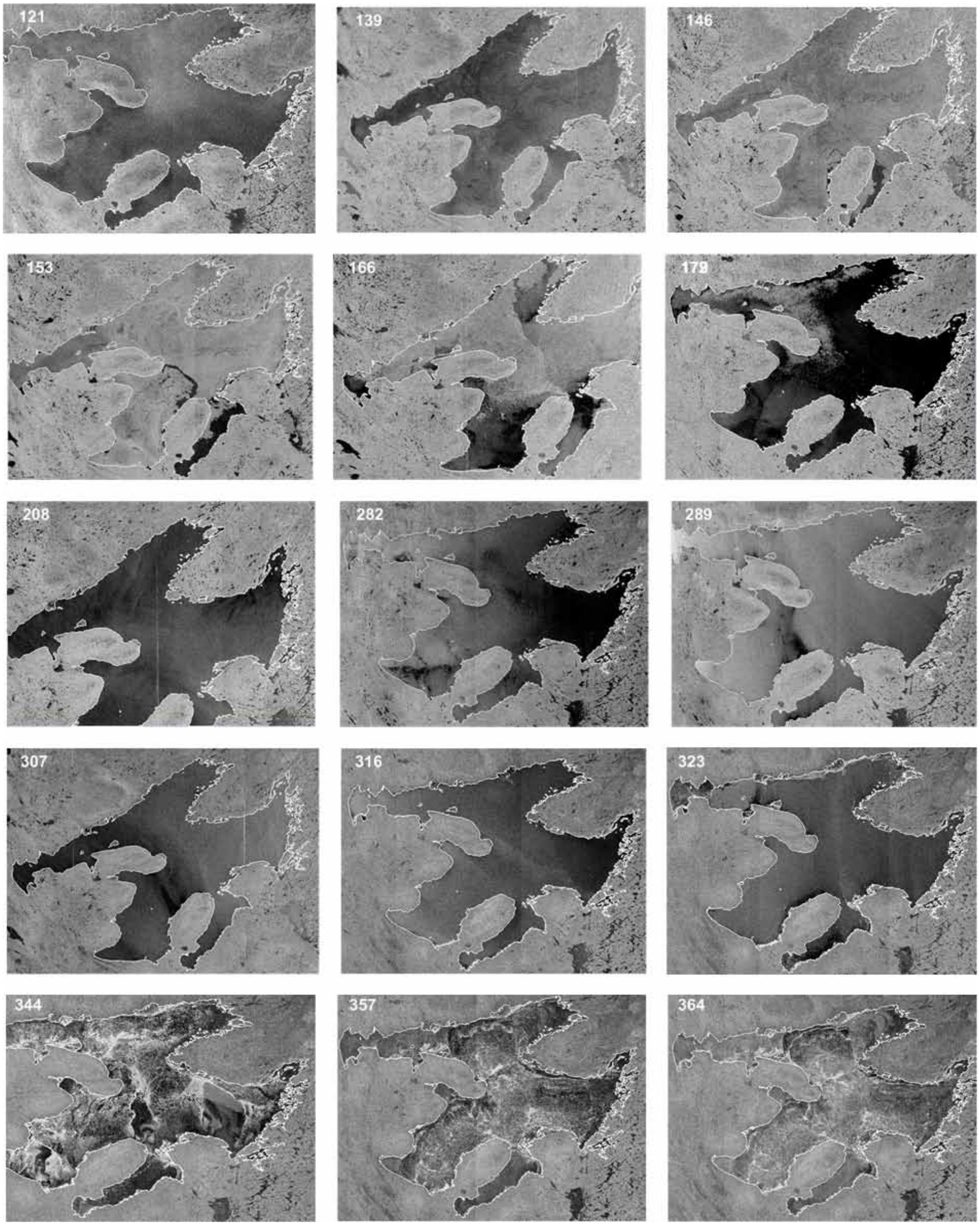


Fig. 6.10: Temporal original SAR images of GBL 1998.

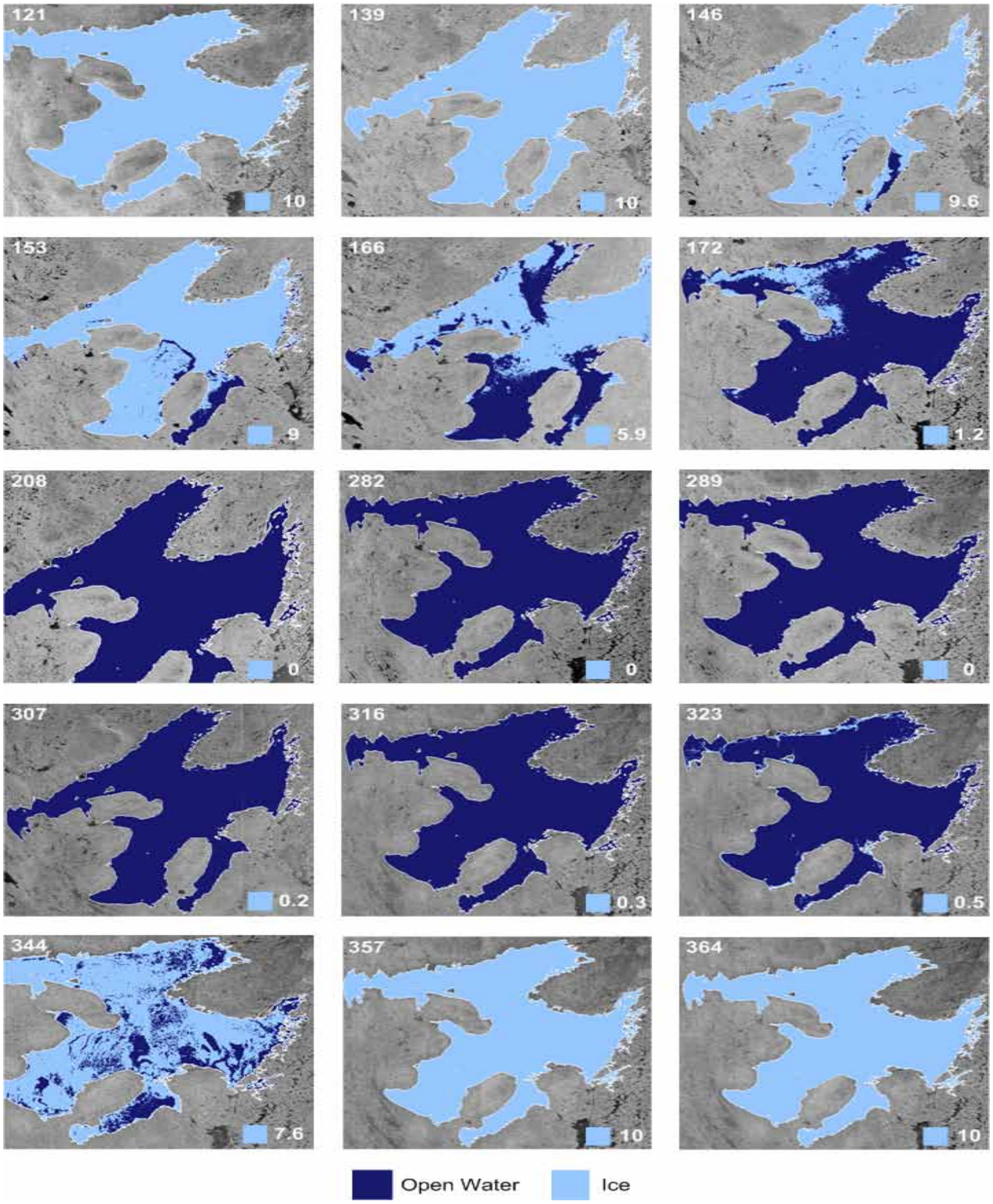


Fig. 6.11: Classification of the images of Fig.6.10 with unsupervised approach. The images cover the cycle of lake ice phenology events. The first image is of early May before the melting season starts. The image of day 139 refers to melt onset with little open water seen in the left bottom arm of the lake. The melting continues from east to west throughout the season until the water clear of ice is reached a day 208. The freeze-up season starts with image of day 307 and the ice gradually covers the lake completely by the date 357 and thickens at later date.

The few entries "na" refer to the dates CIS did not provide ice fraction records. The entries are consistent for all other days. Surprisingly, the unsupervised IRGS approach and CIS ice fraction estimates for many dates closely match each other, but in some days the figures considerably differ with a deviation of almost ±2. The closeness of unsupervised approach results to CIS figures already indicate that the performance of unsupervised approach is within the valid range. To better analyze the deviation in estimates, the results from Table 6.5 historical analysis are graphically plotted in Fig. 6.12. The plots show the evolution of the unsupervised approach and CIS ice fraction estimates versus Julian days and weekly mean temperatures. Noticeable is the smooth transitions in unsupervised approach curve in the graph of GBL 1998, whereas, the CIS curve has abrupt transitions in several days. For instance in day 306 CIS reported the ice fraction change from 0 to "2" which stayed the same for the week of 316 day, the next week the lower ice fraction of "1" has been reported which indicate the melt. However, the detailed analysis of the original images in Fig 6.10 and the steady decrease in temperature profiles in the plot reveals that there is no evidence for the melt. On the other hand the unsupervised approach provides the smoother and realistic transition for both melting and freezing seasons. The same smoothness can be seen in all the graphs of Fig. 6.12.

Table 6.5: Seasonal ice fraction estimates GBL 1997-2007 of unsupervised approach and CIS.

| Year | Day | Unsupervised approach | CIS | Temp | % lake |
|---|---|---|---|---|---|
| 1997 | 191 | 6.6 | 7 | 12.0 | 1 |
| | 283 | 0 | 0 | -5.2 | 0.89 |
| 1998 | 121 | 10 | 10 | -1.8 | 0.99 |
| | 139 | 10 | 9 | 6.7 | 1 |
| | 146 | 9.6 | 10 | 7.3 | 1 |
| | 153 | 9 | 9 | 9.0 | 1 |
| | 166 | 5.9 | na | 8.7 | 0.95 |
| | 172 | 1.2 | na | 9.3 | 1 |
| | 208 | 0 | 0 | 16.6 | 0.82 |
| | 282 | 0 | 0 | -0.2 | 1 |
| | 289 | 0 | 0 | -1.8 | 0.99 |
| | 307 | 0.2 | 2 | -6.3 | 1 |
| | 316 | 0.3 | 2 | -9.6 | 1 |
| | 323 | 0.5 | 1 | -17.8 | 1 |
| | 344 | 7.6 | 9 | -13.6 | 0.95 |
| | 357 | 10 | 10 | -22.1 | 1 |
| | 364 | 10 | 10 | -27.6 | 1 |
| 1999 | 147 | 10 | 10 | 4.1 | 0.94 |
| | 153 | 9.6 | 9 | 7.5 | 0.99 |
| | 274 | 0 | 0 | -2.6 | 0.81 |
| | 291 | 0 | 0 | -5.4 | 0.94 |
| | 328 | 2.3 | 3 | -17.2 | 0.97 |
| | 330 | 6.3 | 9 | -15.8 | 0.91 |
| | 336 | 8.3 | 9 | -22.4 | 1 |
| 2000 | 167 | 9.4 | 9 | 13.2 | 1 |
| | 174 | 8 | 8 | 11.6 | 0.97 |
| | 181 | 5.3 | 7 | 14.2 | 0.92 |
| | 335 | 7.9 | 8 | -17.4 | 0.85 |
| 2001 | 123 | 10 | 10 | -6.7 | 1 |
| | 157 | 10 | 10 | 7.3 | 1 |
| | 171 | 9.5 | 9 | 10.5 | 1 |
| | 180 | 8.9 | 9 | 7.1 | 1 |
| | 187 | 8.3 | 8 | 12.4 | 1 |
| | 192 | 6.5 | 7 | 11.1 | 0.97 |
| | 319 | 0.7 | 2 | -11.6 | 0.92 |
| | 326 | 1.4 | 2 | -10.0 | 0.78 |
| | 332 | 6.1 | 9 | -21.5 | 1 |
| | 333 | 7.1 | 9 | -22.7 | 0.51 |
| | 339 | 10 | 10 | -22.3 | 1 |
| | 346 | 10 | 10 | -27.2 | 0.98 |
| | 353 | 10 | 10 | -26.7 | 1 |
| 2002 | 129 | 10 | 10 | -7.1 | 0.78 |
| | 135 | 10 | 10 | 4.1 | 1 |
| | 142 | 10 | 10 | 2.3 | 1 |
| | 149 | 10 | 10 | 1.0 | 1 |
| | 156 | 10 | 10 | 9.5 | 1 |
| | 170 | 9.5 | 9 | 7.4 | 0.43 |
| | 178 | 9.3 | 8 | 12.8 | 0.89 |
| | 179 | 8.9 | 8 | 9.3 | 0.95 |
| | 185 | 6.4 | 7 | 13.7 | 0.99 |
| | 192 | 1.3 | 1 | 11.6 | 1 |
| | 326 | 3.9 | 5 | -13.3 | 1 |
| 2003 | 170 | 9.1 | 9 | 9.5 | 1 |
| | 184 | 4.9 | 3 | 11.1 | 1 |
| | 191 | 2.7 | 1 | 13.2 | 0.99 |
| 2004 | 170 | 10 | 10 | 9.8 | 0.95 |
| | 177 | 9.7 | 9 | 12.9 | 0.97 |
| | 184 | 9.3 | 8 | 11.1 | 0.92 |
| | 190 | 8.9 | 7 | 13.7 | 1 |
| | 197 | 6.7 | 6 | 10.9 | 1 |
| | 204 | 4.8 | 2 | 10.0 | 1 |
| 2005 | 167 | 9.5 | 9 | 10.6 | 1 |
| | 174 | 8.6 | 8 | 7.4 | 1 |
| | 181 | 7.1 | 7 | 10.6 | 1 |
| | 188 | 3.6 | 4 | 10.8 | 1 |
| | 202 | 0 | 0 | 10.1 | 0.68 |
| | 208 | 0 | 0 | 8.7 | 1 |
| | 216 | 0 | 0 | 10.3 | 0.54 |
| | 280 | 0 | 0 | 0.4 | 0.45 |
| | 287 | 0 | 0 | -0.9 | 1 |
| | 294 | 0 | 0 | -3.8 | 1 |
| | 300 | 0 | 0 | -2.4 | 1 |
| | 335 | 8.3 | 9 | -17.5 | 1 |
| | 349 | 8.6 | 9 | -13.5 | 1 |
| 2006 | 27 | 10 | 10 | -31.9 | 1 |
| | 34 | 10 | 10 | -31.1 | 1 |
| | 152 | 9.8 | 9 | 4.6 | 1 |
| | 159 | 9.2 | na | 8.1 | 0.95 |
| | 166 | 6.4 | 7 | 16.0 | 0.97 |
| | 173 | 1.3 | 3 | 14.0 | 0.67 |
| | 186 | 0 | 0 | 11.2 | 1 |
| | 193 | 0 | 0 | 11.0 | 1 |
| | 200 | 0 | 0 | 15.2 | 0.99 |
| | 207 | 0 | 0 | 15.5 | 0.95 |
| | 251 | 0 | 0 | 12.0 | 1 |
| | 265 | 0 | 0 | 8.3 | 1 |
| | 285 | 0 | 0 | 0.1 | 0.70 |
| | 292 | 0 | 0 | -3.7 | 0.94 |
| | 333 | 10 | 10 | -27.2 | 0.68 |
| | 340 | 10 | 10 | -22.4 | 0.95 |
| | 358 | 10 | 10 | -16.6 | 0.97 |
| 2007 | 122 | 10 | 10 | -5.9 | 0.47 |
| | 136 | 10 | 10 | -0.2 | 0.84 |
| | 173 | 9.3 | 9 | 8.4 | 1 |
| | 178 | 8 | 8 | 13.1 | 1 |
| | 284 | 0 | 0 | -4.0 | 1 |
| | 291 | 0 | 0 | -5.4 | 1 |
| | 301 | 0 | 0 | -9.3 | 1 |
| | 310 | 1 | 2 | -9.4 | 1 |
| | 327 | 9 | 7 | -14.6 | 0.96 |
| | 339 | 10 | 10 | -24.4 | 0.99 |
| | 346 | 10 | 10 | -20.3 | 1 |
| | 353 | 10 | 10 | -18.3 | 0.97 |
| | 359 | 10 | 10 | -24.6 | 0.91 |

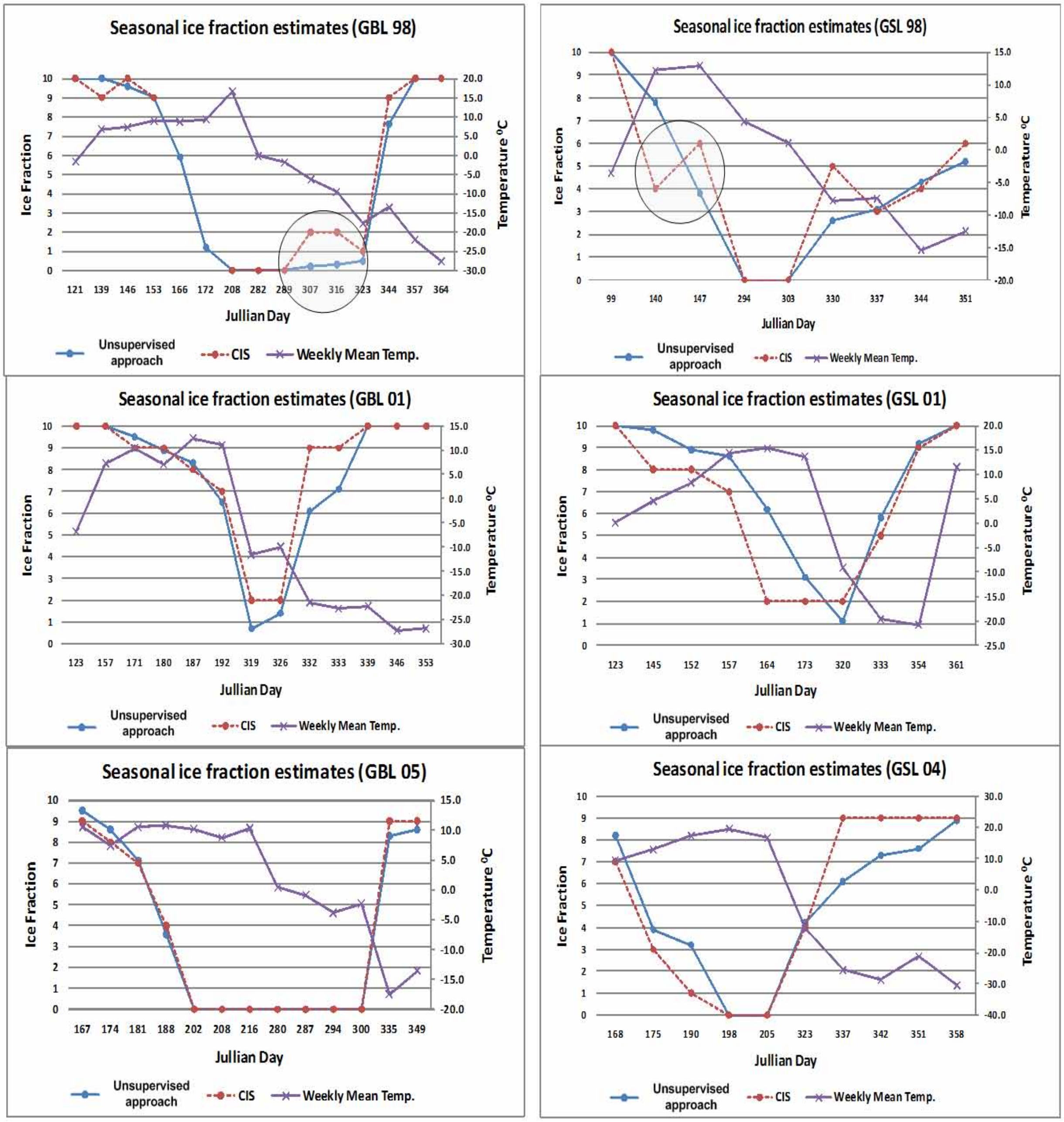


Fig. 6.12: Comparative graph of seasonal ice fractions estimated by unsupervised approach and CIS for the GBL 1998/2001/2005 and GSL 1998/2001/2004.

### Great Slave Lake

An additional experiment was performed with GSL images in Fig. 6.13. The results in Fig 6.14 and Table. 6.6 are presented as an application of unsupervised approach to a different lake. Fewer images are available for GSL. The performance of unsupervised approach remains consistent even in the challenging day 337 with rough water in the freeze-up period. In days 344 and 351 of 1998 (Fig. 6.14) there are misclassifications in the arms of the lakes. The reason is that the open water has a different characteristic in the arms than in the middle of the lake for this scenario. Ice fraction estimates for year 1998 are plotted in Fig. 6.12 from unsupervised classification and CIS. The maximum difference in values is almost 3.7 at day 140 for which the unsupervised classification is successful (Fig 6.13, 6.14). The same abrupt change as in GBL 1998 (Section 6.5.2) is seen for day 147 where the CIS figure indicates that there is an increase in ice fraction as much as ”2” from the previous week, whereas, the temperature continued to increase and the images in Figs. 6.13, 6.14 clearly indicate further melt as opposed to freezing.

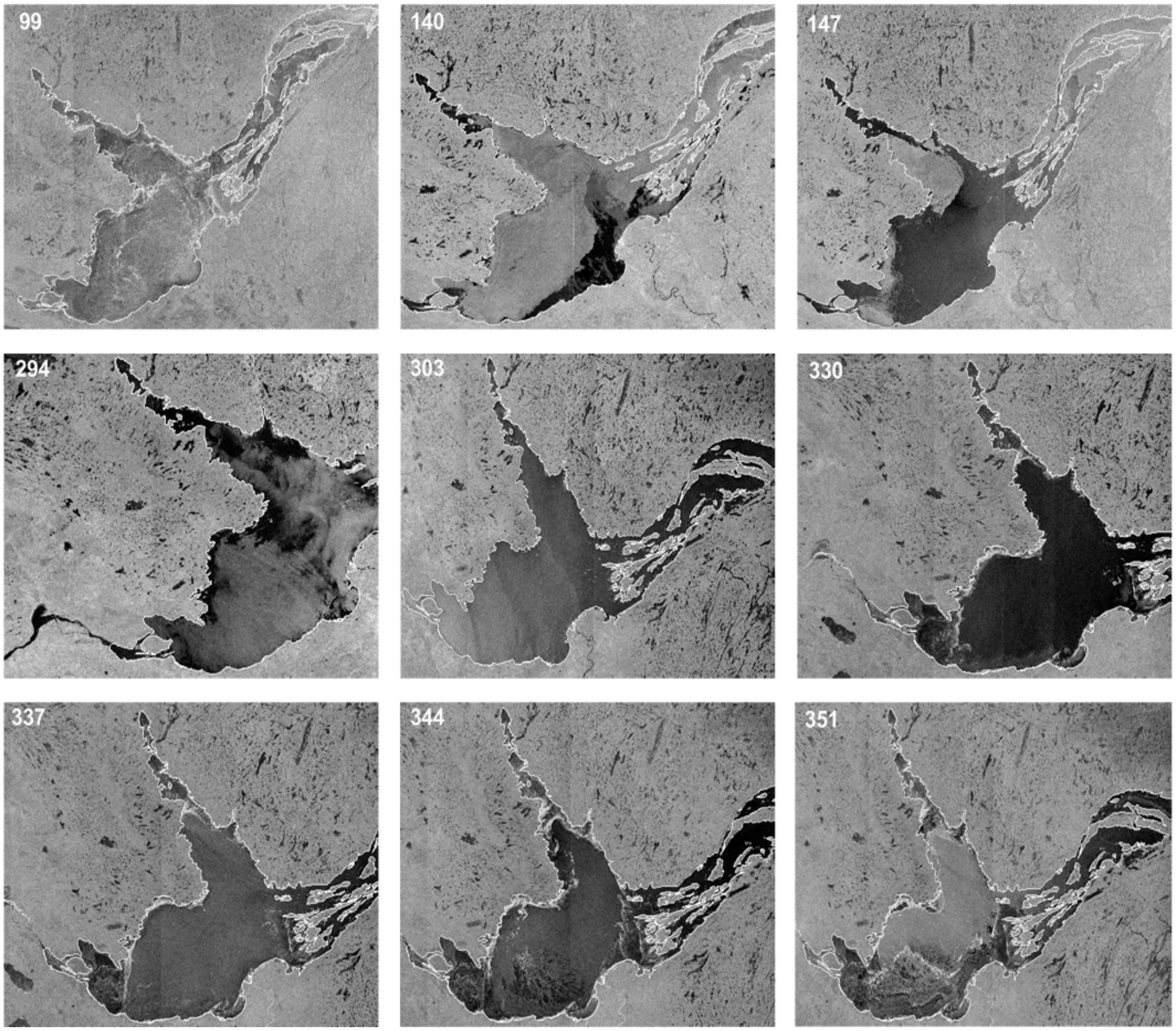


Fig. 6.13: Temporal original SAR images of GSL 1998.

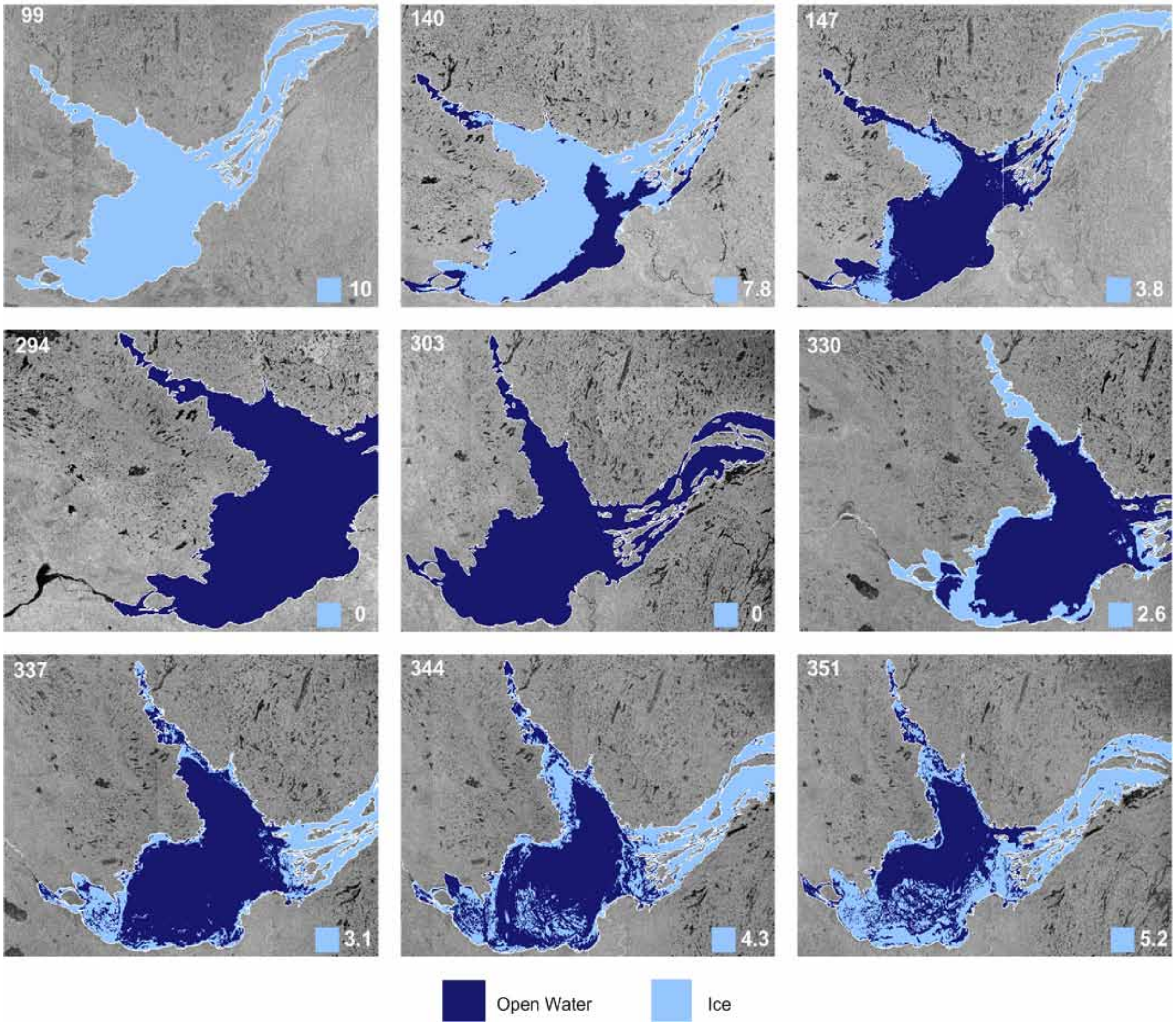


Fig. 6.14: Classification of the images of Fig.6.13 with unsupervised approach. The ice is consolidated in the image of day 99. The lake starts melting from southern part of lake and gradually expands to northern parts of the lake and arms. The results with unsupervised approach are capturing well the changes from melt. The example of the unsupervised approach performance in freeze up period is demonstrated on day 330. Although, the ice in the body of the lake was correctly located by the unsupervised approach the arms have been misclassified as ice. The reason is the rough water in the middle which is different from the water in the arms.

Table 6.6: Seasonal ice fraction estimates GSL 1997-2007 of unsupervised approach and CIS.

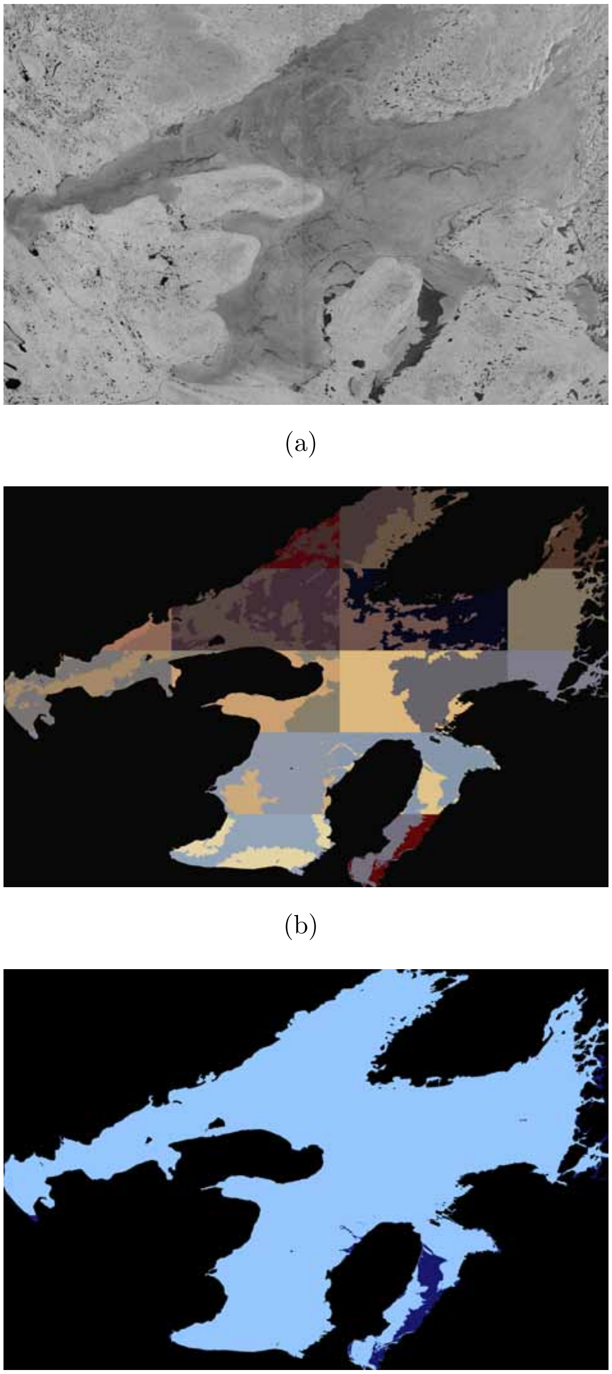
| Year | Day | Unsupervised approach | CIS | Temp | % lake |
|---|---|---|---|---|---|
| 1997 | 308 | 0 | 0 | -5.6 | 0.88 |
| | 317 | 1.8 | 2 | -4.5 | 0.87 |
| | 323 | 2 | 2 | -11.2 | 1 |
| 1998 | 99 | 10 | 10 | -3.5 | 1 |
| | 140 | 7.8 | 4 | 12.2 | 0.96 |
| | 147 | 3.8 | 6 | 13.0 | 1 |
| | 294 | 0 | 0 | 4.3 | 0.73 |
| | 303 | 0 | 0 | 1.1 | 0.98 |
| | 330 | 2.6 | 5 | -7.8 | 0.78 |
| | 337 | 3.1 | 3 | -7.5 | 0.85 |
| | 344 | 4.3 | 4 | -15.4 | 0.93 |
| | 351 | 5.2 | 6 | -12.5 | 0.98 |
| 1999 | 7 | 10 | 10 | -31.7 | 1 |
| | 21 | 10 | 10 | -19.2 | 0.56 |
| | 146 | 9.2 | 7 | 9.1 | 0.92 |
| | 153 | 7.5 | na | 9.2 | 0.68 |
| | 161 | 2.6 | 1 | 8.5 | 1 |
| | 301 | 0.9 | 1 | -0.7 | 0.78 |
| | 308 | 1.7 | 1 | -5.4 | 0.78 |
| | 315 | 2 | 1 | -5.5 | 0.93 |
| | 358 | 10 | 10 | -21.3 | 1 |
| 2000 | 158 | 3.3 | 3 | 7.1 | 1 |
| | 175 | 2 | 2 | 16.3 | 0.79 |
| | 322 | 2.7 | 2 | -7.8 | 0.95 |
| | 329 | 3.1 | 2 | -10.3 | 0.59 |
| 2001 | 123 | 10 | 10 | 0.2 | 0.45 |
| | 145 | 9.8 | 8 | 4.6 | 1 |
| | 152 | 8.9 | 8 | 8.4 | 1 |
| | 157 | 8.6 | 7 | 14.4 | 0.99 |
| | 164 | 6.2 | 2 | 15.4 | 0.89 |
| | 173 | 3.1 | 2 | 13.6 | 0.38 |
| | 320 | 1.1 | 2 | -9.1 | 1 |
| | 333 | 5.8 | 5 | -19.6 | 0.86 |
| | 354 | 9.2 | 9 | -20.8 | 1 |
| | 361 | 10 | 10 | 11.6 | 1 |
| 2002 | 3 | 10 | 10 | -15.5 | 1 |
| | 123 | 10 | 10 | 5.1 | 1 |
| | 129 | 10 | 10 | 14.0 | 0.74 |
| | 136 | 10 | 10 | 7.5 | 0.88 |
| | 157 | 9.3 | 9 | 11.9 | 1 |
| | 164 | 8.1 | 7 | 9.7 | 1 |
| | 171 | 4 | 4 | 12.1 | 1 |
| | 179 | 1.3 | 2 | 15.7 | 0.92 |
| | 185 | 0.5 | 2 | 13.9 | 0.69 |
| | 326 | 3.6 | 3 | -14.8 | 0.68 |
| 2003 | 162 | 2.7 | 3 | 11.2 | 0.95 |
| | 184 | 0 | 0 | 14.5 | 0.78 |
| | 190 | 0 | 0 | 16.1 | 0.86 |
| | 337 | 2.8 | 6 | -19.4 | 0.90 |
| 2004 | 168 | 8.2 | 7 | 9.4 | 0.93 |
| | 175 | 3.9 | 3 | 13.0 | 0.70 |
| | 190 | 3.2 | 1 | 17.3 | 0.98 |
| | 198 | 0 | 0 | 19.4 | 0.78 |
| | 205 | 0 | 0 | 16.8 | 0.82 |
| | 323 | 4.2 | 4 | -12.3 | 0.78 |
| | 337 | 6.1 | 9 | -25.6 | 0.89 |
| | 342 | 7.3 | 9 | -28.7 | 0.76 |
| | 351 | 7.6 | 9 | -21.2 | 0.82 |
| | 358 | 8.9 | 9 | -30.5 | 0.67 |
| 2005 | 146 | 8.8 | 9 | 3.3 | 0.92 |
| | 160 | 4.9 | 4 | 14.3 | 1 |
| | 203 | 0 | 0 | 13.8 | 1 |
| | 208 | 0 | 0 | 11.3 | 1 |
| | 280 | 0 | 0 | 1.3 | 0.92 |
| | 287 | 0 | 0 | 2.3 | 0.96 |
| | 294 | 0 | 0 | -2.9 | 0.84 |
| | 300 | 0 | 0 | 2.3 | 0.85 |
| | 307 | 0.8 | 1 | -3.0 | 0.92 |
| | 323 | 1.2 | 3 | -15.0 | 1 |
| | 343 | 9.4 | 9 | -16.1 | 0.95 |
| 2006 | 6 | 9 | 9 | -12.5 | 1 |
| | 19 | 10 | 10 | -24.3 | 0.72 |
| | 26 | 10 | 10 | -22.8 | 0.82 |
| | 116 | 10 | 10 | 1.2 | 0.70 |
| | 146 | 3.9 | 4 | 8.0 | 0.85 |
| | 150 | 1.9 | na | 7.9 | 1 |
| | 160 | 1.1 | 1 | 11.0 | 1 |
| | 167 | 0 | 0 | 18.0 | 0.52 |
| | 174 | 0 | 0 | 18.4 | 0.70 |
| | 187 | 0 | 0 | 15.9 | 0.85 |
| | 208 | 0 | 0 | 17.3 | 1 |
| | 294 | 0.8 | 1 | -2.6 | 1 |
| | 307 | 1.2 | 1 | -8.8 | 0.84 |
| | 321 | 2 | 4 | -11.9 | 0.90 |
| | 356 | 10 | 10 | -14.5 | 0.70 |
| | 362 | 10 | 10 | -13.5 | 0.85 |
| 2007 | 100 | 10 | 10 | -3.7 | 0.92 |
| | 109 | 10 | 10 | -3.4 | 0.95 |
| | 116 | 10 | 10 | 0.6 | 0.84 |
| | 123 | 10 | 10 | 4.4 | 0.60 |
| | 143 | 9.1 | 9 | 3.1 | 0.100 |
| | 150 | 9.1 | 8 | 10.1 | 0.100 |
| | 172 | 0.9 | 2 | 11.1 | 0.90 |
| | 333 | 6.4 | 5 | -23.1 | 0.95 |

### 6.5.3 Bootstrap approach

The manual labeling performed in Section 6.5.2 is an exhaustive exercise. The integrator is involved at many stages: determining and providing number of classes prior to unsupervised classification, merging and labeling the regions. The implication has become more apparent when processing a large image database. To solve this problem, classification has been performed with a bootstrap approach and the results have been compared with images obtained from the blind test. As mentioned in Section 3.5, the purpose is to evaluate the bootstrap approach with reference images obtained without *a priori* knowledge of unsupervised classification results.

#### Polygon data

The same methodology as in Section 4.1 was used to generate the polygons. The lake surface is partitioned into polygons and the labels (ice/open water) are assigned. The polygons are depicted in Fig. 6.15(b), Fig. 6.16(b), Fig. 6.17(b) and Fig. 6.18(b). The rule of thumb is: the more complex the scene is the more polygons are required for successful classification. The images in Fig. 6.15 and Fig. 6.17 correspond to the melting season where contrast between the classes is high (consolidated ice distinctive from from open water). In such a scenario, a few polygons would be sufficient. For example 17 and 14 polygons are used for classification of images in Fig.6.15 and Fig.6.17, respectively. More polygons, 42, are chosen for the challenging scene in Fig. 6.16 of freezing season. The concentration of ice in this scene is low, therefore, more polygons have been utilized to achieve a mix of classes. The number of polygons can be reduced if the ones with only open water are merged. In the middle of the scene there are polygons containing only open water. Such an approach is discussed in Section 4.1. The lake in Fig. 6.18 has open water with varying backscatter characteristics. Rough water in the center and calm water in the arms makes the classification of lake challenging. Hence, more polygons have been utilized as shown in Fig. 6.18(b).

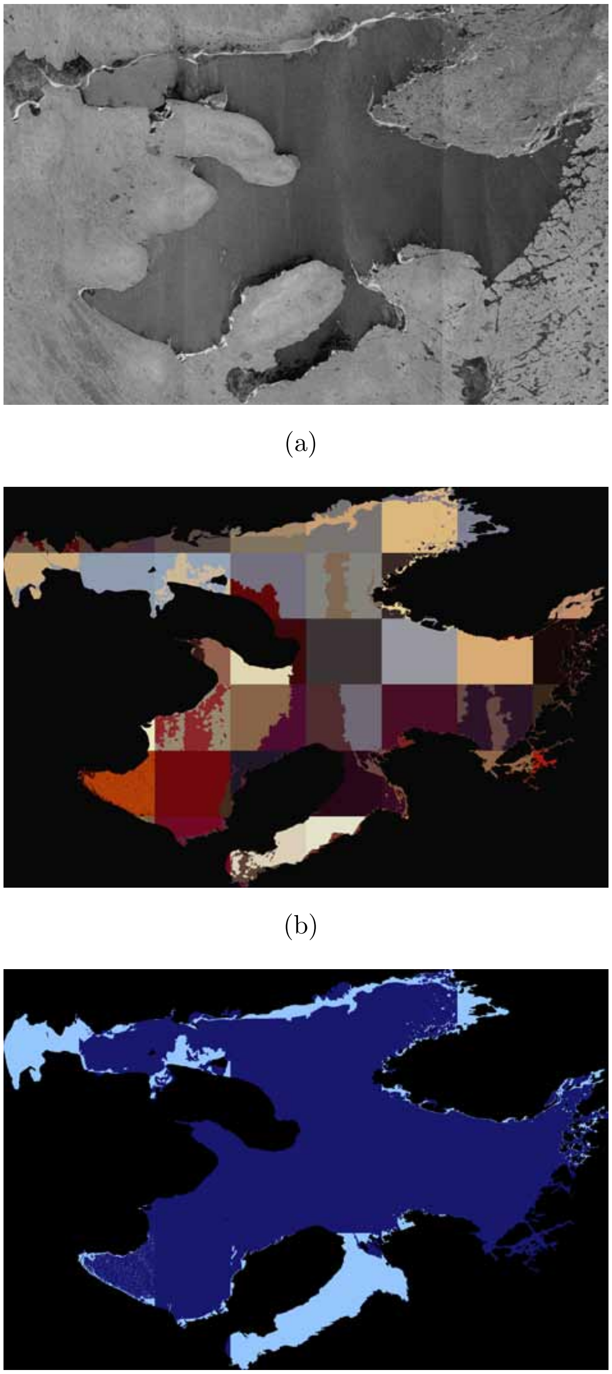


(a)

(b)

(c)

Fig. 6.15: Classification with bootstrap approach. (a) Original image, GBL, May 26 1998. (b) Polygon generation (17 polygons) and unsupervised classification. (c) Automatic labeling result $accuracy = 98.4\%$ , $kappa = 0.76$.

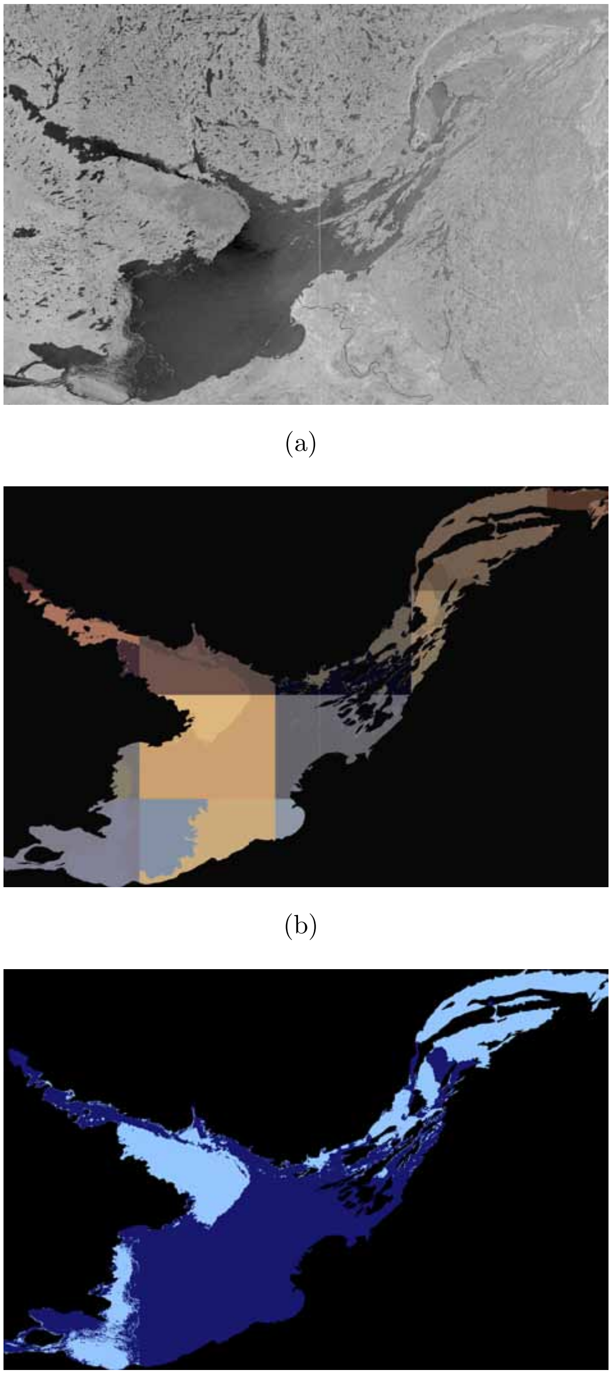


(a)

(b)

(c)

Fig. 6.16: Classification with bootstrap approach. (a) Original image, GBL, Nov 19 1998. (b) Polygon generation (42 polygons) and unsupervised classification. (c) Automatic labeling result $accuracy = 89.2\%$ , $kappa = 0.62$.



(a)

(b)

(c)

Fig. 6.17: Classification with bootstrap approach. (a) Original image, GSL, May 27 1998. (b) Polygon generation (14 polygons) and unsupervised classification. (c) Automatic labeling result $accuracy = 94.1\%$ , $kappa = 0.87$.

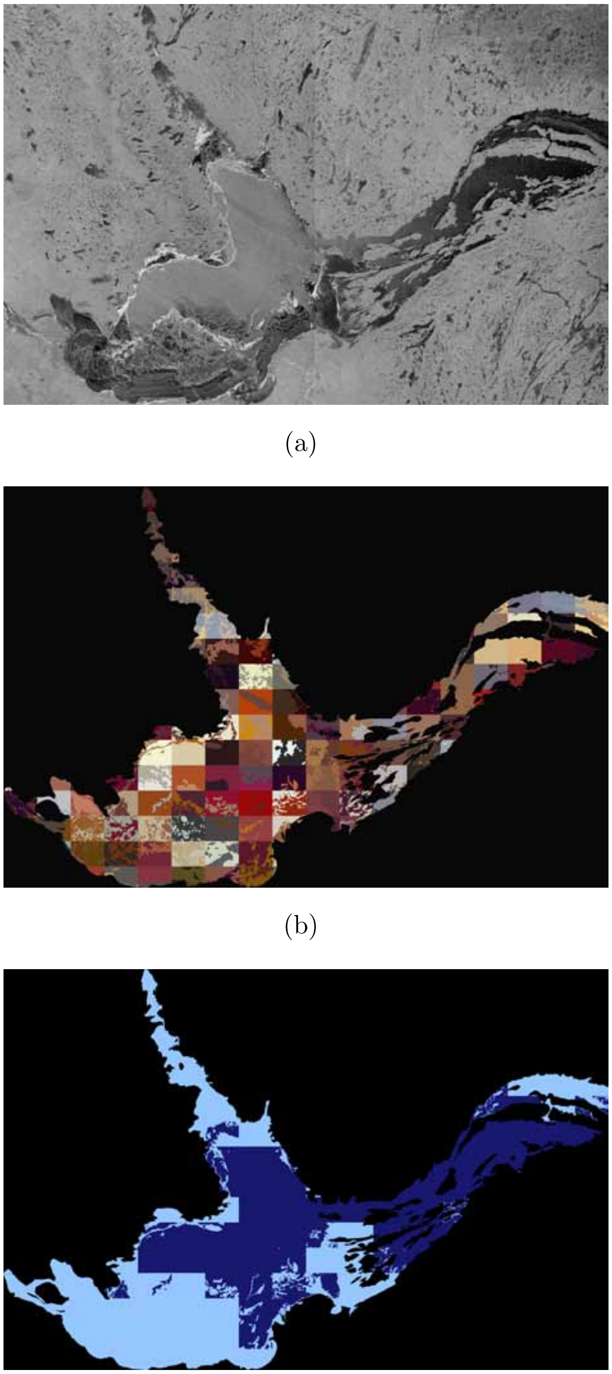


Fig. 6.18: Classification with bootstrap approach. (a) Original image, GSL, Dec 17 1998. (b) Polygon generation (121 polygons) and unsupervised classification. (c) Automatic labeling result $accuracy = 86.3\%$ , $kappa = 0.73$.

### Lake ice polygon unsupervised classification

During the investigation of unsupervised classification methodologies in Section 6.5.1, IRGS [2] demonstrated a better performance, therefore was selected as the preferred unsupervised classification technique for the bootstrap approach. Unsupervised classification was performed on all generated polygons as shown in Figs. 6.15, 6.15, 6.15, 6.18(b) where the produced regions overlayed with polygon data.

### Lake ice image labeling

Following the unsupervised classification, the automatic labeling was performed as part of the bootstrap approach. Figs. 6.15, 6.15, 6.15, 6.18(c) depict the final outputs of classification with the bootstrap approach where the pixels of lake are labeled as ice/no ice. The ice and open water have been successfully combined into classes in images of melting season, GBL Fig. 6.15(c) and GSL Fig. 6.17(c) which validates the concept of the bootstrap approach. As seen from Table 6.7, the performance of the bootstrap approach is high for a higher number of polygons. The tradeoff is setting the initial labeling constraints which becomes a fast exercise if the number of polygons is less, therefore, for images Fig. 6.15 and Fig. 6.17 the number of polygons can be chosen as 17 and 14 or even less. The general rule for polygon generation is described in earlier in this section. In Fig. 6.16(a) the ice is newly formed and scattered. Compared to the whole scene, the ice fraction is low and it is intuitive to capture ice in the selected polygons. The results in Table 6.7 for GBL image of November 19, indicate that if a higher number of polygons is chosen the ice and open water are better separated as depicted in the classification output of Fig. 6.16(a) with 42 polygons (Fig. 6.16(b)). However, in this particular image (Fig. 6.16(a)), the performance is acceptable even with a fewer number of polygons. The requirement for selecting more polygons is evident in the difficult scene in Fig. 6.18(a)) where the inter and intra class variability is high. As presented in Table 6.7, the performance significantly decreases as the number of polygons decreases in difficult scenes such as GSL, December 17. If the polygons are selected as described in Section 6.5.3, the polygon number should be set to better capture local characteristics of classes. The classification output in Fig. 6.18(c) depicts consistent labeling of ice and open water where 121 polygons were used (Fig. 6.18(b)). As the number of polygons increases, the block artifacts become more apparent.

Table 6.7: Performance (*accuracy*/*kappa*) of bootstrap approach as a function of number of polygons.

GBL May 26

| Number of polygons | Performance (*accuracy*/*kappa*) |
|---|---|
| 37 | 98.7%/0.81 |
| 17 | 98.4%/0.76 |
| 10 | 98.6%/0.79 |
| 6 | 98.1%/0.69 |

GBL Nov 19

| Number of polygons | Performance (*accuracy*/*kappa*) |
|---|---|
| 42 | 89.9%/0.65 |
| 13 | 83.6%/0.44 |
| 8 | 86.5%/0.54 |
| 4 | 83.2%/0.35 |

GSL May 27

| Number of polygons | Performance (*accuracy*/*kappa*) |
|---|---|
| 37 | 95.1%/0.89 |
| 14 | 94.1%/0.87 |
| 8 | 94.8%/0.88 |
| 5 | 94.3%/0.87 |

GSL Dec 17

| Number of polygons | Performance (*accuracy*/*kappa*) |
|---|---|
| 121 | 86.3%/0.73 |
| 44 | 59.4%/0.17 |
| 8 | 55.6%/0.10 |
| 5 | 38.2%/ − 0.23 |

### Comparison of classification approaches

The results summarized in Table 6.8 compare unsupervised and the bootstrap approach for image classification. The results indicate the high accuracy and kappa coefficient of classification with the bootstrap approach. Especially, the significant improvement have been achieved classifying challenging images of freezing dates using the bootstrap approach.

Table 6.8: Comparison of classification approaches(unsupervised approach with manual labeling) and bootstrap approach. All reference images are generated by blind test. GBL/GSL 1998 (accuracy/kappa). Number of generated polygons is 17, 42, 14, 121, respectively (Table 6.7)

| Classification approach | GBL May 26 | GBL Nov 19 | GSL May 27 | GSL Dec 17 |
|---|---|---|---|---|
| Unsupervised | 97.9%/0.74 | 72.4%/0.28 | 89.3%/0.77 | 64.6%/0.29 |
| Bootstrap | 98.4%/0.76 | 89.2%/0.62 | 94.1%/0.87 | 86.3%/0.73 |

# Chapter 7

# Summary and Future Directions

A new look into image classification has been introduced with the bootstrap approach. The bootstrap approach is designed to be generic and the application areas of the technique are broad. The summary of the technique has been presented. The classification of difficult scenes such as SAR sea-ice and lake ice images are examples of real life applications of bootstrap approach in an operational environment. The operational considerations and future directions are given at the end.

## 7.1 Summary of Contributions

### 7.1.1 Bootstrap approach for image classification

In the bootstrap approach the image is divided into polygons with constrained set of labels automatically or manually. Next, unsupervised classification performed on each polygon, independently. Finally, using the joint information from all polygons the optimal configuration of labels is found based on objective function derived from Markov random field (MRF) model. The outcome is a generic, fast, repeatable and accurate methodology that generates fully classified images. The technique has been verified with scenes of nature and artificial image. Overall, new perspective has been introduced for image classification.

### 7.1.2 Bootstrap approach for sea-ice classification

A bootstrap approach has been designed and implemented for operational SAR sea-ice image classification. In the classification process, the SAR sea-ice images are classified utilizing provided polygon data. The bootstrap approach for sea-ice classification 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.

The concept has been demonstrated with operational SAR images provided by CIS. 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.

### 7.1.3 Bootstrap approach for lake ice classification

The potential of using automatic unsupervised classification algorithms for SAR lake ice interpretation has been explored on lakes GBL/GSL. The regions produced by IRGS better reflected the separation of different classes compared to other unsupervised classification techniques. However, in difficult scenes of freezing period the unsupervised classification of a full scene with IRGS has yielded lower performance.

The manual labeling with MAGIC has been used for ice fraction estimation and its results have been compared with CIS. The experiment on eleven years of temporal data has shown that the MAGIC and CIS results are mostly consistent. In cases when CIS reported abrupt changes MAGIC provided smoother transitions better describing the lake ice dynamics.

The SAR lake ice interpretation process becomes fully automated if the labeling process of ice and open water is also automated which has been achieved with the bootstrap approach. Bootstrap approach produced a high accuracy classifying both the images of melting and freezing season compared to traditional classification approaches.

### 7.1.4 Methodology for performance evaluation

All reference images used for performance evaluation of bootstrap approach are provided by third party experts. The procedures followed to obtain accurate and unbiased evaluation of developed algorithm is a key contribution. The novel performance evaluation framework has been developed to validate the unsupervised classification and labeling of SAR sea-ice images. Systematic arms length evaluation has been conducted with a trained sea-ice expert using this framework to obtain a set of reference images used for comparison.

## 7.2 Future Directions

### 7.2.1 Operational considerations

Until now the proposed method has demonstrated high performance over all real and synthetic images but the reader should be aware of limitations that might affect the performance. There are three main limitations that might occur:

1. Inaccurate polygon contours. Could happen if human generates polygons. To minimize the error the user should err towards setting polygon boundaries separating the region with the same class. Model for bootstrap approach is generic and handles both ways of polygon generation.

2. Inaccurate unsupervised classification. The technique selected for unsupervised classification should be the best for application. Additionally, the polygons should be selected to avoid having underrepresented classes which are statistically insignificant for unsupervised classification techniques.

3. The repetition of labels and the continuity of polygons is an important factor affecting the labeling performance and uniqueness of a solution. The polygons should be generated to constrain the labeling solution and the generation of isolated polygons should be avoided.

The temporal resolution of SAR images used in this thesis is coarse and weekly based, however, with the constellation of satellites the more frequent data is expected to be available in near future. The current automation scheme has already shown the contribution by producing fully classified maps of sea-ice and analysis of lake events in terms of ice fractions. The bootstrap approach presented with examples in this thesis can be applied to other sea and lakes for operational monitoring.

### 7.2.2 Future research

The basis for classification with bootstrap approach has been established. Modeling regions as MRF and labeling with probabilistic framework along with implementation scheme is a reliable foundation for further development. The current implementation makes use of tone as a feature, the edge penalty as a spatial context, as a domain level knowledge it uses the set of constrained labels provided for each polygon. All this information does not require training and is available within the image. However, the method can be extended to multivariate features and can integrate additional domain knowledge specific to application. For instance, the incidence angle, wind, snow cover, speckle noise and many other factors are forcing the exploration of features different from tone to shed the light about the characteristics of the ice-type. The sea-ice experts themselves use the ancillary information such as the state of previous satellite overpass and also make a judgement with regard to the ice-type existing in specific region based on shape and textural properties. Recently, with the launch of RADARSAT 2 the cross co-polarization bands are also contributing in decisions by providing a better separation between open water and ice and the better emphasis on the locations of thick ice. Thus, the research is foreseen in the areas of texture, shape feature extraction from the regions and the best way of utilization of cross-co polarization band. Applying bootstrap approach to other areas is also foreseen.

# References

[1] http://www.wmo.int/. World Meteorological Organization website [online], March 2009. x, 4, 13, 14, 15, 65, 69, 70

[2] Q. Yu and D. A. Clausi. IRGS: Image segmentation using edge penalties and region growing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(12):2126–2139, 2008. xiii, 5, 8, 15, 18, 21, 22, 26, 27, 28, 41, 68, 105

[3] DA Clausi, AK Qin, MS Chowdhury, P. Yu, and P. Maillard. Magic: Map-guided ice classification system. *Canadian Journal of Remote Sensing*, 36(1):183–194, 2009. xiii, 15, 26, 59, 60

[4] R. Hoffman and A.K. Jain. Segmentation and classification of range images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (5):608–620, 2009. 1

[5] R.C. Gonzales and R.E. Woods. Digital Image Processing. *Prentice-Hall*, 2002. 1, 40

[6] C. Samson, L. Blanc-Féraud, G. Aubert, and J. Zerubia. A level set model for image classification. *International Journal of Computer Vision*, 40(3):187–197, 2000. 1

[7] A.K. Jain, R.P.W. Duin, and J. Mao. Statistical pattern recognition: A review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(1):4–37, 2000. 1, 2

[8] S.D. Olabarriaga and AWM Smeulders. Interaction in the segmentation of medical images: A survey. *Medical image analysis*, 5(2):127–142, 2001. 1, 2

[9] P. Bernad, L. Denise, and P. Refregier. Hierarchical feature-based classification approach for fast and user-interactive SAR image interpretation. *Geoscience and Remote Sensing Letters, IEEE*, 6(1):117–121, 2009. 1, 2, 18, 21

[10] D. Lu and Q. Weng. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5):823–870, 2007. 1, 2, 24

[11] R.K.S. Kwan, A.C. Evans, and G.B. Pike. MRI simulation-based evaluation of image-processing and classification methods. *Medical Imaging, IEEE Transactions on*, 18(11):1085–1097, 1999. 2

[12] E.L. Chen, P.C. Chung, C.L. Chen, H.M. Tsai, and C.I. Chang. An automatic diagnostic system for CT liver image classification. *Biomedical Engineering, IEEE Transactions on*, 45(6):783–794, 1998. 2

[13] S.V. Aschkenasy, C. Jansen, R. Osterwalder, A. Linka, M. Unser, S. Marsch, and P. Hunziker. Unsupervised image classification of medical ultrasound data by multiresolution elastic registration. *Ultrasound in medicine & biology*, 32(7):1047–1054, 2006. 2

[14] S. Lee and M.M. Crawford. Unsupervised multistage image classification using hierarchical clustering with a Bayesian similarity measure. *Image Processing, IEEE Transactions on*, 14(3):312–320, 2005. 2

[15] G. Camps-Valls, B. Marsheva, and D. Zhou. Semi-supervised graph-based hyperspectral image classification. *Geoscience and Remote Sensing, IEEE Transactions on*, 45(10):3044–3054, 2007. 2

[16] M. Mignotte, C. Collet, P. Perez, and P. Bouthemy. Sonar image segmentation using an unsupervised hierarchical MRF model. *Image Processing, IEEE Transactions on*, 9(7):1216–1231, 2000. 2

[17] S. Ochilov and D. A. Clausi. Automated classification of operational SAR sea ice images. In *Computer and Robot Vision (CRV), 2010 Canadian Conference on*, pages 40–46. IEEE, 2010. 2, 5, 8, 24

[18] S. Ochilov, NA Svacina, CR Duguay, and DA Clausi. Towards an automated lake ice monitoring system from sar imagery. In *AGU Fall Meeting Abstracts*, volume 1, 2010. 2

[19] RA English, SJ Rawlinson, and NM Sandirasegaram. *Development of an ATR Workbench for SAR Imagery.* Citeseer, 2002. 2

[20] G. Gao. Statistical Modeling of SAR Images: A Survey. *Sensors*, 10(1):775–795, 2010. 2

[21] C. Tison, J.M. Nicolas, F. Tupin, and H. Maître. A new statistical model for Markovian classification of urban areas in high-resolution SAR images. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(10):2046–2057, 2004. 2

[22] R. Fjortoft, Y. Delignon, W. Pieczynski, M. Sigelle, and F. Tupin. Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(3):675–686, 2003. 2

[23] F. Galland, N. Bertaux, and P. Réfrégier. Minimum description length synthetic aperture radar image segmentation. *Image Processing, IEEE Transactions on*, 12(9):995–1006, 2003. 2, 18, 21

[24] Y. Wu, K. Ji, W. Yu, and Y. Su. Region-based classification of polarimetric SAR images using Wishart MRF. *Geoscience and Remote Sensing Letters, IEEE*, 5(4):668–672, 2008. 2, 21

[25] D. Dai, W. Yang, and H. Sun. Multilevel Local Pattern Histogram for SAR Image Classification. *Geoscience and Remote Sensing Letters, IEEE*, 8(2):225–229, 2011. 2

[26] R. Latifovic and D. Pouliot. Analysis of climate change impacts on lake ice phenology in Canada using the historical satellite data record. *Remote Sensing of Environment*, 106(4):492–507, 2007. 3

[27] J.J. Magnuson, D.M. Robertson, B.J. Benson, R.H. Wynne, D.M. Livingstone, T. Arai, R.A. Assel, R.G. Barry, V. Card, E. Kuusisto, et al. Historical trends in lake and river ice cover in the northern hemisphere. *Science*, 289(5485):1743, 2000. 3

[28] W.L. Anderson, D.M. Robertson, and J.J. Magnuson. Evidence of climatic change and projected future change. *Limnology and Oceanography*, 41(5):815–821, 1996. 3

[29] D.K. Hall, D.B. Fagre, F. Klasner, G. Linebaugh, and G.E. Liston. Analysis of ERS 1 synthetic aperture radar data of frozen lakes in northern Montana and implications for climate studies. *Journal of Geophysical Research-Oceans*, 99(C11), 1994. 3, 76

[30] C.R. Duguay, T.D. Prowse, B.R. Bonsal, R.D. Brown, M.P. Lacroix, and P. Ménard. Recent trends in Canadian lake ice cover. *Hydrological Processes*, 20(4):781–801, 2006. 3

[31] D.M. Livingstone. Break-up dates of Alpine lakes as proxy data for local and regional mean surface air temperatures. *Climatic Change*, 37(2):407–439, 1997. 3

[32] C.R. Duguay, G.M. Flato, M.O. Jeffries, P. Ménard, K. Morris, and W.R. Rouse. Ice-cover variability on shallow lakes at high latitudes: model simulations and observations. *Hydrological Processes*, 17(17):3465–3483, 2003. 3, 74, 76

[33] Q. Yu and D.A. Clausi. Sar sea-ice image analysis based on iterative region growing using semantics. *Geoscience and Remote Sensing, IEEE Transactions on*, 45(12):3919–3931, 2007. 5, 8, 17, 28, 29

[34] R. Nock and F. Nielsen. Grouping with bias revisited. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages 460–465. 6, 27

[35] P.K. Mallapragada, R. Jin, and A.K. Jain. Active query selection for semi-supervised clustering. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2009. 6, 27

[36] AK Jain, P.K. Mallapragada, and M. Law. Bayesian feedback in data clustering. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 3, pages 374–378. IEEE, 2006. 6, 27

[37] R. Adams and L. Bischof. Seeded region growing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(6):641–647, 2002. 6

[38] D. Geman, S. Geman, C. Graffigne, and P. Dong. Boundary detection by constrained optimization. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(7):609–628, 2005. 6

[39] Y. Liu, R. Jin, and A.K. Jain. Boostcluster: boosting clustering by pairwise constraints. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 450–459. ACM, 2007. 6

[40] S.X. Yu and J. Shi. Grouping with bias. In *Proc. of NIPS*, pages 1327–1334, 2001. 6, 27

[41] T. Lange, M.H.C. Law, A.K. Jain, and J.M. Buhmann. Learning with constrained and unlabelled data. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. Proceedings of the 2005 IEEE Computer Society Conference on*, volume 1, pages 731–738. IEEE Computer Society, 2005. 6, 27

[42] S.X. Yu and J. Shi. Segmentation given partial grouping constraints. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(2):173–183, 2004. 6, 27

[43] J. Wang, E. Gu, and M. Betke. MosaicShape: stochastic region grouping with shape prior. 1:902–908, 2005. 6

[44] Z. Lu and T. Leen. Semi-supervised learning with penalized probabilistic clustering. *Advances in neural information processing systems*, 17:849–856, 2005. 6

[45] L. Zhang and Q. Ji. A bayesian network model for automatic and interactive image segmentation. *Image Processing, IEEE Transactions on*, (99):1–1, 2011. 6, 23

[46] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, 1986. 7

[47] D.G. Barber and E.F. LeDrew. SAR sea ice discrimination using texture statistics: A multivariate approach. *Photogrammetric Engineering and Remote Sensing*, 57(4):385–395, 1991. 8, 17, 24

[48] A.V. Bogdanov, S. Sandven, O.M. Johannessen, V.Yu. Alexandrov, and L.P. Bobylev. Multisensor approach to automated classification of sea ice image data. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(7):1648–1664, 2005. 8, 17, 24

[49] D. A. Clausi. Comparison and fusion of co-occurrence, Gabor and MRF Texture Features for Classification of SAR Sea-Ice Imagery. *Atmosphere and Oceans*, 39(3):183–194, 2001. 8, 17, 24

[50] D. Haverkamp, LK 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*, volume 2, pages 638–640. IEEE, 1993. 8, 24, 25

[51] M.E. Shokr. Evaluation of second-order texture parameters for sea ice classification from radar images. *Journal of Geophysical Research*, 96(C6):10625–10, 1991. 8, 17, 24

[52] J.C. Curlander and R.N. McDonough. Synthetic Aperture Radar: Systems and Processing. *John Wiley and Sons, New York*, 1991. 11

[53] http://sv.gina.alaska.edu. Geographic information network of alaska. Swath viewer website [online], June 2010. 11

[54] K. Morris, M. O. Jeffries, and W. F. Weeks. Ice processes and growth history on Arctic and sub-Arctic lakes using ERS-1 SAR data. *Polar Record*, 31(177):115–128, 1995. 11, 74, 75

[55] W.F. Weeks M.O. Jeffries, K. Morris and H. Wakabayashi. Structural and stratigraphic features and ERS 1 synthetic aperture radar backscatter characteristics of ice growing on shallow lakes in NW Alaska, winter 1991-1992. *Journal of Geophysical Research*, 99:22459–22472, 1994. 11, 76

[56] P.M. Lafleur C.R. Duguay, T.J Pultz and D. Drai. RADARSAT backscatter characteristics of ice growing on shallow sub-Arctic lakes, Churchill, Manitoba, Canada. *Hydrological Processes*, 16(8):1631–1644, 2002. 11, 76

[57] C. R. Duguay and P. M. Lafleur. Determining depth and ice thickness of shallow sub-Arctic lakes using space-borne optical and SAR data. *International Journal of Remote Sensing*, 24(3):475–489, 2003. 11

[58] KK Kang, CR Duguay, SEL Howell, CP Derksen, and REJ Kelly. Potential of AMSR-E for Estimating Lake Ice Thickness on Great Bear and Great Slave Lakes, Canada. In *Geophysical Research Abstracts*, volume 10, 2008. 11

[59] S.E.L. Howell, L.C. Brown, K.K. Kang, and C.R. Duguay. Variability in ice phenology on Great Bear Lake and Great Slave Lake, Northwest Territories, Canada, from SeaWinds/QuikSCAT: 2000-2006. *Remote Sensing of Environment*, 113(4):816–834, 2009. 11

[60] http://weblogs.sun sentinel.com/. Quikscat satellite dies website [online], Nov 2009. 11

[61] http://nsidc.org/. National snow and ice data center website [online], Nov 2010. 11

[62] http://ice glaces.ec.gc.ca/. Canadian Ice Service website [online], March 2009. 12

[63] G. Flett Dean, J. Wilson Katherine, W. Vachon Paris, and J. Fred 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*, 28(3):490–497, 2002. 12

[64] D. Lubin and R. Massom. *Polar Remote Sensing: Atmosphere and Oceans*, volume 1. Springer Verlag, 2006. 12

[65] M. Arkett, D. Flett, R. De Abreu, P. Clemente-Colon, J. Woods, and B. Melchior. Evaluating alos-palsar for ice monitoring-what can l-band do for the north american ice service? In *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, volume 5, pages V–188. IEEE, 2008. 12

[66] http://www.geobase.com/. National hydro network(nhn) website [online], Feb 2010. 14

[67] J.P. Snyder. *Map projections–a working manual.* USGPO, 1987. 16

[68] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical recipes in C.* Cambridge university press Cambridge, 1992. 16

[69] C. Oliver and S. Quegan. Understanding Synthetic Aperture Radar Images. *Artech House Publishers, Boston, London*, 1998. 16

[70] D. A. Clausi. Remote Sensing Systems. *Course Notes, Winter*, 2007. 16

[71] L. Soh, C. Tsatsoulis, D. Gineris, and C. Bertoia. ARKTOS: An Intelligent System for SAR Sea Ice Image Classification. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(1):229–248, 2004. 17, 25, 26, 29

[72] M. Petrou and P.G. Sevilla. *Dealing with Texture.* John Wiley & Sons Inc., 2006. 17, 20, 33

[73] P. Maillard, D. A. Clausi, and H. Deng. Map-guided sea ice segmentation and classification using SAR imagery and a MRF segmentation scheme. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(12):2940–2951, 2005. 17, 29

[74] Q. Yu. *Automated SAR Sea Ice Interpretation.* PhD thesis, 2006. 17, 19, 36

[75] L.K. Soh and C. Tsatsoulis. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *Geoscience and Remote Sensing, IEEE Transactions on*, 37(2):780–795, 1999. 17

[76] D. A. Clausi. *Texture Segmentation of SAR Sea Ice Imagery.* PhD thesis, 1996. 17

[77] D. A. Clausi and M.Ed Jernigan. Designing Gabor Filters for optimal texture separability. *Pattern Recognition*, 33(11):1835–1849, 2000. 17

[78] L. Vincent and P. Soille. Watershed in digital spaces: An efficient algorithm based on immersion simulations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(6):583–598, 1991. 17, 18

[79] J. Goodman. Statistical properties of laser speckle patterns. *Laser speckle and related phenomena*, pages 9–75, 1975. 18

[80] EA Carvalho, DM Ushizima, FNS Medeiros, CIO Martins, RCP Marques, and INS Oliveira. SAR imagery segmentation by statistical region growing and hierarchical merging. *Digital Signal Processing*, 20(5):1365–1378, 2010. 18, 21

[81] J.S. Lee. Digital image enhancement and noise filtering by use of local statistics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2):165–168, 1980. 18

[82] V.S. Frost, J.A. Stiles, KS Shanmugan, and J.C. Holtzman. A model for radar images and its application to adaptive digital filtering of multiplicative noise. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2):157–166, 1982. 18

[83] D.T. Kuan, A.A. Sawchuk, T.C. Strand, and P. Chavel. Adaptive noise smoothing filter for images with signal-dependent noise. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2):165–177, 1985. 18

[84] Y. Gauthier, M. Bernier, and J.P. Fortin. Aspect and incidence angle sensitivity in ERS-1 SAR data. *International Journal of Remote Sensing*, 19(10):2001–2006, 1998. 18

[85] M.W. Lang, P.A. Townsend, and E.S. Kasischke. Influence of incidence angle on detecting flooded forests using C-HH synthetic aperture radar data. *Remote Sensing of Environment*, 112(10):3898–3907, 2008. 18

[86] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern Classification.* Citeseer, 2001. 19, 20

[87] W. Yang, C. He, Y. Cao, H. Sun, and X. Xu. Improved Classification of SAR Sea Ice Imagery Based on Segmentation. In *IEEE International Conference on Geoscience and Remote Sensing Symposium IGARSS*, pages 3727–3730, 2006. 21

[88] S.Z. Li. *Markov Random Field Modeling in Computer Vision.* Springer, New York, 2009. 21, 31, 32, 35, 40, 41

[89] H. Derin and H. Elliott. Modeling and segmentation of noisy and textured images using Gibbs random fields. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 9(1):39–55, 1987. 21

[90] H. Deng and D. A. Clausi. Unsupervised image segmentation using a simple MRF model with a new implementation scheme. *Pattern Recognition*, 37(12):2323–2335, 2005. 22, 26, 41

[91] X. Descombes, R.D. Morris, J. Zerubia, and M. Berthod. Estimation of Markov Random Field Prior Parameters Using Markov Chain Monte Carlo Maximum Likelihood. *Image Processing, IEEE Transactions on*, 8(7):954–963, 1999. 22

[92] R. Zabih and V. Kolmogorov. Spatially coherent clustering using graph cuts. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II–437. IEEE, 2004. 23

[93] Y.Y. Boykov and M.P. Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 1, pages 105–112. IEEE, 2001. 23

[94] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(11):1222–1239, 2001. 23

[95] V. Kolmogorov and R. Zabin. What energy functions can be minimized via graph cuts? *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(2):147–159, 2004. 23

[96] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for markov random fields with smoothness-based priors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(6):1068–1080, 2008. 23

[97] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11:285–296, 1975. 23, 27

[98] N.V. Lopes, P.A. Mogadouro do Couto, H. Bustince, and P. Melo-Pinto. Automatic histogram threshold using fuzzy measures. *Image Processing, IEEE Transactions on*, 19(1):199–204, 2010. 23

[99] S. Alpert, M. Galun, R. Basri, and A. Brandt. Image segmentation by probabilistic bottom-up aggregation and cue integration. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007. 23

[100] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):603–619, 2002. 23

[101] S. Krinidis and V. Chatzis. A robust fuzzy local information c-means clustering algorithm. *Image Processing, IEEE Transactions on*, 19(5):1328–1337, 2010. 23

[102] M. Mignotte. A label field fusion bayesian model and its penalized maximum rand estimator for image segmentation. *Image Processing, IEEE Transactions on*, 19(6):1610–1624, 2010. 23

[103] http:/rst.gsfc.nasa.govSect1Sect1_16.html. The Remote Sensing Tutorial Image processing and interpretation - Morro Bay, California as the test image [online], December 2010. 24

[104] L. Hubert-Moy, A. Cotonnec, L. Le Du, A. Chardin, and P. Perez. A comparison of parametric classification procedures of remotely sensed data applied on different landscape units. *Remote Sensing of Environment*, 75(2):174–187, 2001. 24

[105] D. Chen and D. Stow. The effect of training strategies on supervised classification at different spatial resolutions. *Photogrammetric Engineering and Remote Sensing*, 68(11):1155–1162, 2002. 24

[106] D.A. Landgrebe. *Signal theory methods in multispectral remote sensing.* Wiley-Interscience, 2003. 24

[107] P.M. Mather. *Computer processing of remotely sensed images: an introduction.* Wiley, 2004. 24

[108] C.B. Fetterer and J. Ye. Multi-year ice concentration from RADARSAT. In *IEEE International Geoscience and Remote Sensing Symposium (IGRRS'97)*, pages 402–404, 1997. 25

[109] S.V. Nghiem and G.A. Leshkevich. Satellite SAR remote sensing of great lakes ice cover, part 1. Ice backscatter signatures at C band. *Journal of Great Lakes Research*, 33(4):722–735, 2007. 25, 76

[110] G.A. Leshkevich and S.V. Nghiem. Satellite SAR remote sensing of Great Lakes ice cover, part 2. Ice classification and mapping. *Journal of Great Lakes Research*, 33(4):736–750, 2007. 25

[111] H. Deng and D. A. Clausi. Unsupervised segmentation of synthetic aperture radar sea ice imagery using a novel Markov random field model. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(3):528–538, 2005. 26

[112] R. Samadani. A finite mixtures algorithm for finding proportions in SAR Images. *Image Processing, IEEE Transactions on*, 4(8):1182–1186, 1995. 26

[113] 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*, 20(15):3087–3109, 1999. 26

[114] Y. Gauthier, F. Weber, S. Savary, M. Jasek, L.M. Paquet, and M. Bernier. A combined classification scheme to characterise river ice from SAR data. *EARSeL eProceedings*, 5(1):77–88, 2006. 27

[115] B. Tso and P.M. Mather. *Classification Methods for Remotely Sensed Data.* Taylor & Francis, New York, 2009. 27

[116] R.M. Haralick, K. Shanmugam, and I. Dinstein. Texture features for image classification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 3(6):610–621, 1973. 27

[117] Anil K. Jain and Richard C. Dubes. Algorithms for Clustering Data. *Prentice Hall College Div*, 1988. 27

[118] X.S. Zhou and T.S. Huang. Relevance feedback in image retrieval: A comprehensive review. *Multimedia systems*, 8(6):536–544, 2003. 27

[119] J.W. Modestino and J. Zhang. A markov random field model based approach to image interpretation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(6):606–615, 1992. 28

[120] P. Fieguth. *Statistical Image Processing and Multidimensional Modeling.* Springer, 2010. 32

[121] R.O. Duda, P.E. Hart, and D.G. Stork. Pattern classification. 2001. *NY John Wiley*, 2001. 33, 43

[122] G. Winkler. *Image Analysis, Random Fields and Dynamic Monte Carlo Method.* Springer, New York, 1995. 41

[123] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1091, 1953. 41

[124] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 6(6):721–741, 1984. 41, 43

[125] J. Zhang. The mean field theory in EM procedures for Markov random fields. *IEEE Transactions on Signal Processing*, 40(10):2570–2583, 1992. 41

[126] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, (10):1–38, 1977. 41

[127] G.J. McLachlan T. Krishnan. *The EM Algorithm and Extensions.* Wiley New York, 1997. 41

[128] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. 8th Int'l Conf. Computer Vision*, volume 2, pages 416–423, July 2001. 43, 49

[129] J. Cohen et al. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46, 1960. 48, 77

[130] P.D.N. Hebert and M. McGinley. Great Bear Lake/Great Slave Lake, Northwest Territories. 2008. 72, 73

[131] L. Johnson. Physical and chemical characteristics of Great Bear Lake, Northwest Territories. 1971. 72

[132] http://www.climate.weatheroffice.gc.ca/. National Climate Data and Information Archive website [online], Aug 2010. 73

[133] P.S. Ray. Broadband complex refractive indices of ice and water. *Applied Optics*, 11(8):1836–1844, 1972. 74

[134] K.M. Walter, M. Engram, C.R. Duguay, M.O. Jeffries, and FS Chapin III. The Potential Use of Synthetic Aperture Radar for Estimating Methane Ebullition From Arctic Lakes1. *JAWRA Journal of the American Water Resources Association*, 44(2):305–315, 2008. 76

[135] W.P. Adams and N.T. Roulet. Illustration of the roles of snow in the evolution of the winter cover of a lake. *Arctic*, 33(1):100–116, 1980. 76

[136] R.G. Congalton. A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment*, 37(1):35–46, 1991. 77