Engineering Applications of Artificial Intelligence 25 (2012) 853–868

Contents lists available at SciVerse ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Subglacial water presence classification from polar radar data

Christopher M. Gifford [a,*], Arvin Agah [b,1]

[a] National Security Technology Department, The Johns Hopkins University Applied Physics Laboratory, Laurel, MD 20723, USA
[b] Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS 66045, USA

ARTICLE INFO

Article history:
Received 15 March 2011
Received in revised form
30 August 2011
Accepted 16 December 2011
Available online 4 February 2012

Keywords:
Pattern recognition
Machine learning
Ensemble classification
Radar remote sensing
Subglacial water

ABSTRACT

Ground and airborne radar depth-sounding of the Greenland and Antarctic ice sheets have been used for many years to remotely determine characteristics such as ice thickness, subglacial topography, and mass balance of large bodies of ice. Ice coring efforts have supported these radar data to provide ground truth for validation of the state (wet or frozen) of the interface between the bottom of the ice sheet and the underlying bedrock. Subglacial state governs the friction, flow speed, transport of material, and overall change of the ice sheet. In this paper, we utilize machine learning and classifier combination to model water presence from airborne polar radar data acquired on Greenland in 1999 and 2007. The underlying method results in radar independence, allowing model transfer from 1999 to 2007 radar data to produce water presence maps of the Greenland ice sheet with differing radars. We focus on how to construct a successful set of classifiers capable of high classification accuracy. Utilizing multiple machine learning algorithms is shown to be successful for this classification problem, achieving 86% classification accuracy in the best case. Several heuristics are presented for constructing teams of multiple classifiers for predicting subglacial water presence. The presented methodology could also be applied to radar data acquired over the Antarctic ice sheet.

1. Introduction

Remote sensing methods, such as radar and seismic/acoustic surveys, attempt to acquire data in order to infer properties of the subsurface from a remote location such as the surface, air, or space. A prominent example is the use of radar sensors to gather data from the polar ice sheets, specifically internal layers and the ice–bedrock interface, from the surface of an ice sheet. Other examples are satellite-based imagery and identification of land-cover and events, and the observation of light and its characteristics to infer properties of distant galaxies and stars.

Ground and airborne depth-sounding of the Greenland and Antarctic ice sheets have been used for many years to determine characteristics such as ice thickness, subglacial topography, and mass balance of large bodies of ice. Radar sounding of ice sheets is challenging due to the rough surface interface, various stages of melting both on top of and within the ice sheet, and spatial variation of ice thickness and bedrock topography. Processing the data, including the incorporation of knowledge about the sensing medium, is important for proper interpretation and dissemination of accurate data to the scientific community. For example, in Greenland, water is present over continuous finite distances which are smooth, but do not qualify as large lakes that are evident in Antarctica.

At the University of Kansas, the Center for Remote Sensing of Ice Sheets (CReSIS, 2009) performs polar research to gather data and model ice sheets to better understand global climate changes and the possible effects, including sea level impacts. We have designed, built, and utilized mobile robots to autonomously traverse polar terrain in Greenland and Antarctica and support radar remote sensing data acquisition (Gifford et al., 2009; Stansbury et al., 2004; Harmon et al., 2004; Akers et al., 2004, 2006a, 2006b). Airborne systems have been developed to offer large-scale studies of polar regions. One primary technology for this research is the development of sophisticated radars. CReSIS researchers have provided first views of the ice sheet bed in fast-flowing areas and certain internal ice sheet layers, and continue to provide detailed subsurface images and models describing their behavior and dynamics. CReSIS has a wide variety of radar data from various radar designs over many years. All of these data are available to the public, providing information such as latitude, longitude, radar travel times, bed echo intensity, and ice thickness for extended flight segments. This data repository can then be used for advanced machine learning, data mining, and modeling efforts for polar environments.

In this paper, we utilize ice-penetrating radar data collected in Greenland in May 1999 and September 2007 as part of a model-creation effort for subglacial water presence classification. Using radar

* Corresponding author. Tel.: +1 443 778 4958; fax: +1 443 778 9188.
E-mail addresses: Christopher.Gifford@jhuapl.edu (C.M. Gifford), agah@ku.edu (A. Agah).
[1] Tel.: +1 785 864 8821; fax: +1 785 864 3226.

0952-1976/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2011.12.002

data from ice sheets, the goal is to learn a model of water presence or absence for each radar measurement. This enables the production of predicted water presence maps for the scientific community without needing to drill resource-intensive ice cores. Specifically, a detailed study of ensemble learning and decision combination has been conducted. Experimental results are presented, including successful ensemble compositions, individual learning algorithm contribution, team size, and team diversity. Finally, classification results as well as ensemble learning heuristics discovered as part of this investigation are discussed.

2. Background and related work

Machine learning techniques have seen limited application to ice sheet and polar subsurface data. Most efforts involve identification or tracing of specific layers that hold historical importance. Internal layers of ice sheets have been investigated to predict the depth and thickness of certain layers. For example, initial efforts to predict the depth and thickness of the Eemian Layer in the Greenland ice sheet utilized a Monte Carlo Inversion of the flow model to estimate unknown parameters constrained by the internal layers (Buchardt and Dahl-Jensen, 2007a, 2007b). One of these parameters is the basal melt rate, an important parameter for ice sheet models. Similar work has been done to classify the presence of bottom crevasses, and to estimate their height using radar data of the Ross Ice Shelf in Antarctica (Peters et al., 2007). A physical model was developed by studying the radar data and long echo tails that were found to be characteristic of bottom crevasses. That work largely involved studying the data to determine power reflection coefficients, and did not incorporate any regression or machine learning methods.

Learning classifiers have been recently employed for autonomous polar event detection via satellites. As part of NASA's New Millennium Program, the Autonomous Sciencecraft Experiment (ASE) has been used for detecting dynamic events—such as volcanic eruptions, floods, and cryospheric events—using onboard science algorithms. One such focus was on the detection of ice-based surface events (lakes freezing/thawing and sea ice breakup) (Castano et al., 2005, 2006). As part of that work, four pixel-based classifiers (manually constructed classifier, exhaustive threshold band ratio search classifier, decision tree, and SVM) were employed to generate models to recognize events using hyperspectral images. The SVM classifier, which has seen regular use onboard the EO-1 spacecraft, successfully classified sea ice breakup near Antarctica which autonomously triggered a spacecraft reaction. Such a classifier can be used to identify high priority data and reduce the bandwidth for data communication back to Earth. In Srivastava and Stroeve (2003), preliminary results are presented for unsupervised discovery of geophysical processes (including snow, ice, and clouds) from a spaceborne instrument over Greenland. Their results showed that regional classifiers provide higher classification accuracy for polar regions, but still had difficulty differentiating between certain clouds and snow/ice. Others have taken a knowledge-engineering approach to build an intelligent satellite sea ice classifier by creating rules from sea ice experts (Soh and Tsatsoulis, 2000).

Efforts to classify subsurface layers, occasionally called facies in the literature, using ground penetrating radar (GPR) are abundant. Together with seismic surveying, GPR allows the remote sensing of subsurface properties to guide decisions of where to drill for oil or natural gas, as well as help explain observed surface changes. Radar reflection patterns can be used to distinguish echo returns in different regions, which lends well to machine learning methods to associate such patterns with distinct classes (Moysey et al., 2005). GPR signal interpretation is one of the more studied applications of subsurface classification, as interpretation of subsurface materials or layers requires coring or drilling to determine ground truth. Experts are required to manually study the materials to segment them into distinct rock or object classes. These data can then be used by classifiers to learn a model between GPR return signals and the excavated ground truth. As GPR has been used for many applications, machine learning can aid in automating these tasks by learning a model, interpolating between measurement sites, and characterizing the subsurface at other unknown locations.

From these works, it has been shown that using machine learning to create models utilizing polar radar data analysis is not only feasible, but also can provide high levels of accuracy while offering a significant increase in efficiency. In the following sections, we discuss the approach, using multiple learning algorithms, to further increase accuracy and efficiency for the application of subglacial water presence classification from radar data acquired on the Greenland ice sheet. This approach could also be applied to the Antarctic and other large bodies of ice, but is not explicitly discussed in this paper.

3. Radar, coring, and subglacial water

As part of NASA's Program for Arctic Regional Climate Assessment (PARCA) and others, the University of Kansas and CReSIS have collected ground and airborne radar data in Greenland which cover the majority of the continent's ice sheet. CReSIS now hosts these data for public use. Other major efforts, such as ice coring, have produced data that complement these radar data sets to study subglacial activity of the Greenland ice sheet.

Studying the extent/presence of subglacial water is important, as the state of the bed of the ice sheet (wet or frozen) provides information about flow, friction, and roughness of the underlying bedrock interface. The presence of water (wet state) means that the interface is likely slippery, or that heat has caused a downflow of water from upper or nearby portions of the ice sheet. Portions of the ice with larger amounts of water exhibit faster flow properties and dynamic changes of the ice in those regions due to lubrication at the bed. The lack of water means that the interface is frozen, likely exhibiting more friction and therefore little movement. A smooth bedrock interface lends more to sliding of the ice sheet along its surface, whereas a rough interface introduces more friction between the ice sheet and bedrock on which it rests. If water is present at the bed, then the interface between the ice and water will be smoother, compared to a mixed water and rock interface (which will exhibit the shape of the underlying rock). Transport of material, however, may cause this lower rock interface to change over time.

Ice coring offers an additional method to study, among many other things, the state of the ice–bedrock interface by manually examining the bottom of the cored ice column. Thus, ice cores represent the only ground truth available for large-scale ice sheet modeling and remote sensing validation. In Greenland, there are two primary drill sites which we focus on in this work, namely, GRIP and N-GRIP. The Greenland Ice Core Project (GRIP) drilled a core from 1989 to 1992 to a depth of 3029 m, located at 72.58 N, −37.63 W (NCDC, 2009). The subglacial state of this core was found to be frozen. The North GRIP (N-GRIP) drilled a core from 1999 to 2003 to a depth of 3085 m, located at 75.1 N, −42.32 W (NGRIP, 2009). The subglacial state of this core was found to be wet (high melt).

As a radar transmits energy down to and into the ice—typically from one or more transmitting antennas mounted on the wings of an aircraft—changes in dielectric properties cause the