Machine/Deep Learning Models for Sea Ice Thickness and Surface Roughness Predictions

Research/Project Context

While it has been a popular trend to develop state-of-the-art machine/deep learning architectures for environmental studies, not many research works have focused on viewing the sea ice thickness and surface roughness prediction problems from such an approach. Furthermore, the sea ice studies utilized data with lower spatial coverage collected from satellites such as ICESat and CryoSat missions, which make the models less applicable in the sea ice observation [1] [2]. To mitigate the constraints, this project aims to develop Machine/Deep Learning architectures that train regression models taking advantage of comparably abundant Synthetic Aperture Radar (SAR) data, retrieved from RADARSAT Constellation Mission (RCM), with larger spatial and temporal coverages. The RCM consists of three identical (C-Band) satellites passing Canada's far north four times per day in total. Both pixel-based and object-based algorithms will be explored to determine the feasible method for the sea ice detection problems.

Hypothesis and Purpose

The objective of this project is to develop machine/deep learning architectures to produce sea ice thickness/roughness prediction models using RCM SAR training data and Airborne electromagnetic (AEM) expert data. Proposed machine/deep learning architectures are multi-layer neural network (for pixel-wise approach) and Convolutional Neural Network (CNN) coupled with variational autoencoder (VAE) feature engineering (for object-based approach). It would be ideal to have an accuracy improved from simple threshold-based algorithms using more sophisticated CryoSat radar altimeter data [1]. Therefore, a baseline Root Mean Square Error (RMSE) of 0.9 m is chosen. It is expected that a successful model would produce an RMSE closer to 0.3 m, an accuracy of a random forest model [1]. Such assessment foundation is possible since the expert data used for this project and the compared study are both AEM data. Since there is no sufficient previous work done to be used as a baseline of the sea ice surface roughness prediction problem, an RMSE of existing methods will be calculated to be compared. The K-fold validation technique will be used for reliable evaluations.

Data Requirements and Sources

For supervised learning, labels of sea ice thickness and surface roughness will be generated from AEM surveys collected over the Canadian Arctic Archipelago in 2015, 2016 and 2018. The models will be trained based on the Radarsat-2 dataset in the overlapping location and time range. For each pixel, HH and HV polarization bands data are collected covering 100x100 m^2 area with the ScanSAR beam mode. Four other features processed using such information are included in the data: the HV-HH ratio, HH-HV correlation coefficient within a sliding window, and standard deviations of HH and HV within a sliding window. Different sliding window sizes of 5x5, 7x7 and 9x9 could be explored. Potentially, it might be possible to use the sea ice roughness feature to enhance the thickness prediction accuracy [3].

Benchmarks and Timeline

A GitHub repository for the development of the project will be created. From start to finish, the development of the project will include the creation of Jupyter Notebooks for reproducibility, a Python Package for usability and a comprehensive README file. It is anticipated that the training and testing data pre-processing will be done around mid-October. By the midterm report end of October, the simplest architecture for both problems, sea ice thickness and surface roughness, is expected to be produced. The rest of the project period will be used to accomplish the remaining objectives.

Risk Mitigation Strategy

Due to the collection of AEM data densely occurring following the linear trajectory of a flight, an object-based method can reduce a significant amount of expert data by chunking and averaging the labels into one for an object. Semi-supervised learning is an efficient technique to resolve the problem of lacking expert data [4]. Although it is expected that an object-based approach might perform better due to spatial autocorrelation, Spatial Semi-Supervised Learning (SSSL) can mitigate such a problem of a pixel-based method by taking account of neighbouring inter-pixel relationship to reduce speckle-like noise in the prediction [4]. It is possible that just using the raw feature inputs can decrease the performance of the output model. Therefore, several normalization methods will be explored. Autoencoder techniques can be employed for feature reduction/engineering, which can further enhance performance [5].

References

[1] S. Lee, J. Im, J. Kim, M. Kim, M. Shin, H.-cheol Kim, and L. Quackenbush, "Arctic sea ice thickness estimation from CryoSat-2 satellite data using machine learning-based lead detection," *Remote Sensing*, vol. 8, no. 9, p. 698, 2016.

[2] A. A. Petty, N. T. Kurtz, R. Kwok, T. Markus, and T. A. Neumann, "Winter Arctic sea ice thickness from icesat‐2 Freeboards," *Journal of Geophysical Research: Oceans*, vol. 125, no. 5, 2020.

[3] J. C. Landy, A. A. Petty, M. Tsamados, and J. C. Stroeve, "Sea ice roughness overlooked as a key source of uncertainty in cryosat‐2 ice freeboard retrievals," *Journal of Geophysical Research: Oceans*, vol. 125, no. 5, 2020.

[4] R. R. Vatsavai, S. Shekhar, and T. E. Burk, "An efficient spatial semi-supervised learning algorithm," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 22, no. 6, pp. 427–437, 2007.

[5] N. Su, C. Zhang, Y. Yan, C. Zhao, and Z. Tan, "Sea-Ice image classification for Channel navigation in Polar Application," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018.