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

Summary of Progress / Challenges and Overcoming Strategies

Both sea ice thickness and surface roughness prediction problems were attempted in parallel using the same deep learning architectures. For a strict model assessment, test data are sourced from 2018 data images while training data are extracted from 2015 and 2016 data. All training data were normalized to have a value in the range between 0 and 1. The test data are normalized with the same standard maximum and minimum values in the training dataset. A Python script was developed to create datasets by aggregating expert data within a SAR data raster cell. Data within a raster cell are converted to a vector consisting of averaged expert data with features of HH band, HV band, HV/HH ratio, correlation coefficients within sliding windows (size = 5, 7 and 9), the standard deviation of backscattering values within sliding windows, and averaged incidence angles. The number of expert data points aggregated are recorded per sample. This information is used later for dropping samples that may not reflect actual average thickness or roughness by having too few data points. Samples with less than 50 and 10 expert data points are dropped from roughness and thickness datasets, respectively, considering the AEM data collection intervals and raster cell size (100 m). The numbers of remaining samples are roughly around 10,000 for both datasets.

The first attempts were made using a simple three-layer (hidden size = 30) neural network (NN) architecture. Different feature combinations were experimented with (Tables 1 and 2). For both surface roughness and thickness predictions, best mean absolute errors (MAEs) are produced using HH, HV bands with their depolarization factors (cross-polarization ratio and correlation coefficients between the two bands) and mean incidence angle. However, adding standard deviation of backscattering values to the training process deteriorated the performance to a great extent with unreliable standard deviations in MAEs.

The next step aims to take account of neighbouring cells' information into the training (Tables 3 and 4). A square kernel size of 7x7 is chosen to balance the computational cost and amount of data used for training. The very basic approach of flattening the data linearly into vectors with 49x2=98 (from HH and HV bands) features showed poorer performance than using only the two band features (Tables 1, 2, 3 and 4). It is assumed to have resulted from a high complexity in the feature space. A 3D-convolutional autoencoder (3D-CAE) was developed to resolve the problem. Encoded features can also provide new information that is hard to be found from raw data. Two 7x7 windows with HH and HV bands are stacked into a 3D box (7x7x2) and convoluted into 2x2x1x4 encoded data (feature reduction rate = 83.67%). The performances did not improve much when only the encoded 16 features were used for training (Tables 3 and 4). However, since the feature space dimension complexity is reduced, the advantage of this method is emphasized when concatenated with other features. Observations suggest that combining HH, HV, Depolarization factors and incidence angle features with encoded features maximizes the performance. It outperforms with reliable standard deviation in surface roughness prediction compared with single-cell data results (Tables 1 and 3). Nevertheless, there was no improvement in thickness prediction. Another observation indicates incidence angle as a key feature when used with encoded features. Similar to the single-cell method, standard deviations of HH and HV bands resulted in significant confusion again.

Remaining Timeline and Plans

A few more feature engineering methods can be attempted in the remaining timeframe. Grey-level Co-occurrence Matrix (GLCM) products and SOBEL edge detection output are the candidates. A feature attribution calculation technique should be deployed to explain the ambiguous characteristics of feature combinations in a model. With the promising 3D-CAE results as of the midterm reporting, other feature stack combinations can be considered. Once sufficient roughness prediction results are produced, they can be fed into thickness prediction training with expected performance enhancement. Semi-supervised learning can also experiment if time allows. Such attempts should be finished by early December.

Additional Requirements

One of the remaining methods to be attempted is semi-supervised learning. Initial predictions should be performed on SAR images, which include non-sea-ice areas, such as land. It is hard to collect data not knowing which cells fall into such regions. Having masked images will ease the process.

Tables

Table 1. Roughness prediction results using simple 3-layer NN with different feature combinations.

Features used	MAE		Pearson's r (Prediction vs Expert)	
	Average	Standard Deviation	Average	Standard Deviation
HH and HV bands	68.71	8.07	0.51	0.02
HH, HV and depolarization factors	68.27	5.55	0.50	0.16
HH, HV, depolarization factors and incidence angle	63.42	10.49	0.38	0.10
All features	93.33	89.82	0.42	0.06

Table 2. Thickness prediction results using simple 3-layer NN with different feature combinations.

Features used	MAE		Pearson's r (Prediction vs Expert)	
	Average (meter)	Standard Deviation	Average	Standard Deviation
HH and HV bands	0.63	0.04	0.52	0.49
HH, HV and depolarization factors	0.64	0.03	0.49	0.03
HH, HV, depolarization factors and incidence angle	0.61	0.06	0.57	0.02
All features	2.64	1.79	0.43	0.11

Table 3. Roughness prediction results using neighbouring cell data.

Features used	MAE		Pearson's r (Prediction vs Expert)	
	Average	Standard Deviation	Average	Standard Deviation
Flattened 7x7x2 features	93.28	13.08	0.33	0.04
Encoded features only	98.42	6.77	0.47	0.01
Encoded features + HH and HV	72.55	15.05	0.53	0.05
Encoded features + HH, HV and Depolarization factors	84.66	15.81	0.45	0.04
Encoded features + HH, HV, Depolarization factors and incidence angle	53.71	3.57	0.44	0.05
Encoded features + all features	102.99	78.92	0.34	0.16

Table 4. Thickness prediction results using neighbouring cell data.

Features used	MAE		Pearson's r (Prediction vs Expert)	
	Average (meter)	Standard Deviation	Average	Standard Deviation
Flattened 7x7x2 features	0.84	0.21	0.36	0.14
Encoded features only	0.76	0.15	0.44	0.11
Encoded features + HH and HV	0.67	0.10	0.53	0.07
Encoded features + HH, HV and Depolarization factors	0.79	0.08	0.40	0.07
Encoded features + HH, HV, Depolarization factors and incidence angle	0.64	0.11	0.58	0.01
Encoded features + all features	1.32	0.30	0.20	0.10