



pms73@cam.ac.uk

# A deep active learning framework for the detection of subsurface meltwater lakes on Antarctic ice shelves

Paula Suchantke<sup>1</sup>, Rebecca Dell<sup>1</sup>, Neil Arnold<sup>1</sup>

<sup>1</sup>Scott Polar Research Institute, University of Cambridge, UK



Scott Polar Research Institute



UNIVERSITY OF CAMBRIDGE



## Background

- Meltwater buried in the shallow subsurface of ice shelves is an understudied feature of Antarctic surface and firn layer hydrology
- Continent-wide detection and continued monitoring require **automated methods**, deep learning-based approaches have proven successful in Greenland [1][2]

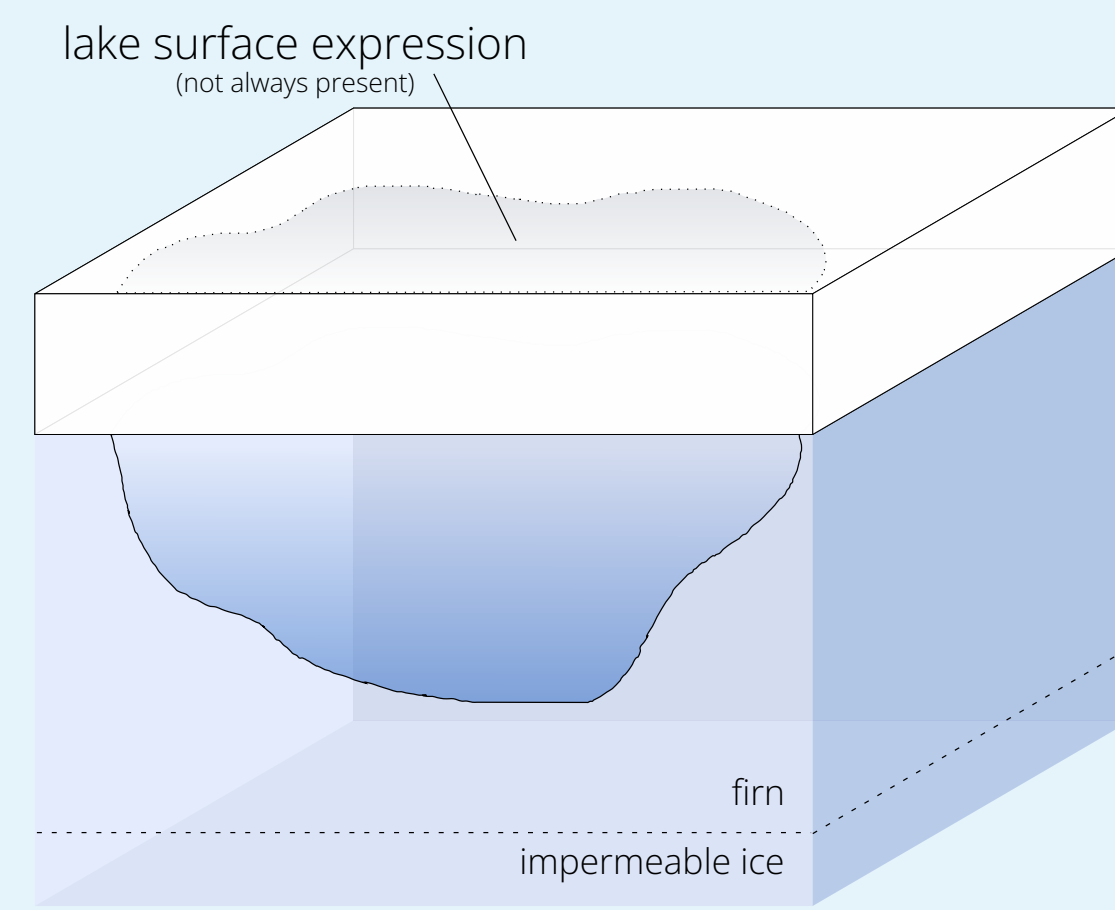
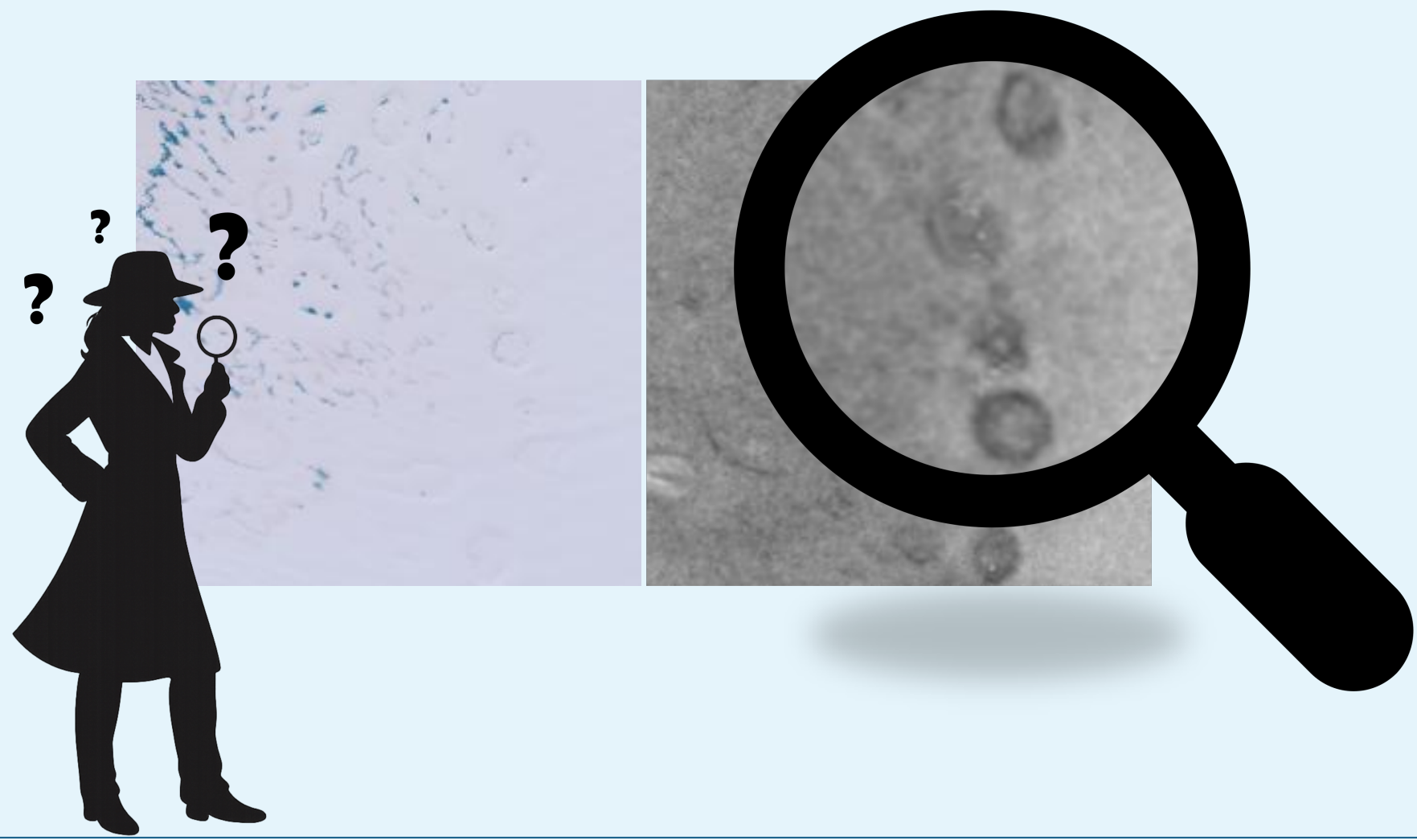


Figure 1: Idealised schematic of a buried lake on an Antarctic ice shelf.

## A Needle in an Ice Stack

- Buried lakes are **not reliably detectable in optical imagery**, necessitating the use of microwave imagery
- They are an **extremely rare class**, requiring methods to address severe class imbalance
- **Ground truth labels are uncertain**, due to ambiguities in interpretation of SAR backscatter



## What's Next?

- Training, testing, and tuning the model
- Applying the model across ice shelves

**Are you an expert on Wilkins Ice Shelf or C-Band SAR interpretation?** Have a look at my timeseries and help me identify these features!



**References:**

[1] Dunmire, D., Barwell, A.F., Wever, N., Lenaerts, J.T., and Datta, R.T., 2021. Contrasting regional variability of buried meltwater extent over 2 years across the Greenland Ice Sheet. *The Cryosphere*, 15(6).

[2] Li, L., Zheng, L., Liang, Q., Cheng, X. and Chen, Z., 2025. Detection of weekly buried meltwater in Greenland Ice Sheet with Sentinel-1 imagery. *International Journal of Digital Earth*, 18(1).

[3] Fischer, G., Papathanassiou, K.P. and Hajnsek, I., 2020. Modeling and compensation of the penetration bias in InSAR DEMs of ice sheets at different frequencies. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13.

[4] Benedek, C.L. and Willis, J.C., 2021. Winter drainage of surface lakes on the Greenland Ice Sheet from Sentinel-1 SAR imagery. *The Cryosphere*, 15(3).

[5] Wang, Y., Braham, N.A.A., Xiong, Z., Liu, C., Albrecht, C.M. and Zhu, X.X., 2023. SSL4EO-S12: A large-scale multimodal, multitemporal dataset for self-supervised learning in Earth observation (Software and Data Sets). *IEEE Geoscience and Remote Sensing Magazine*, 11(3).

[6] Brischler, M., Dietz, A.J., Kneisel, C. and Kuenzer, C., 2021. A novel method for automated supraglacial lake mapping in Antarctica using Sentinel-1 SAR imagery and deep learning. *Remote Sensing*, 13(2).

## Training Data Curation and Labelling Strategy

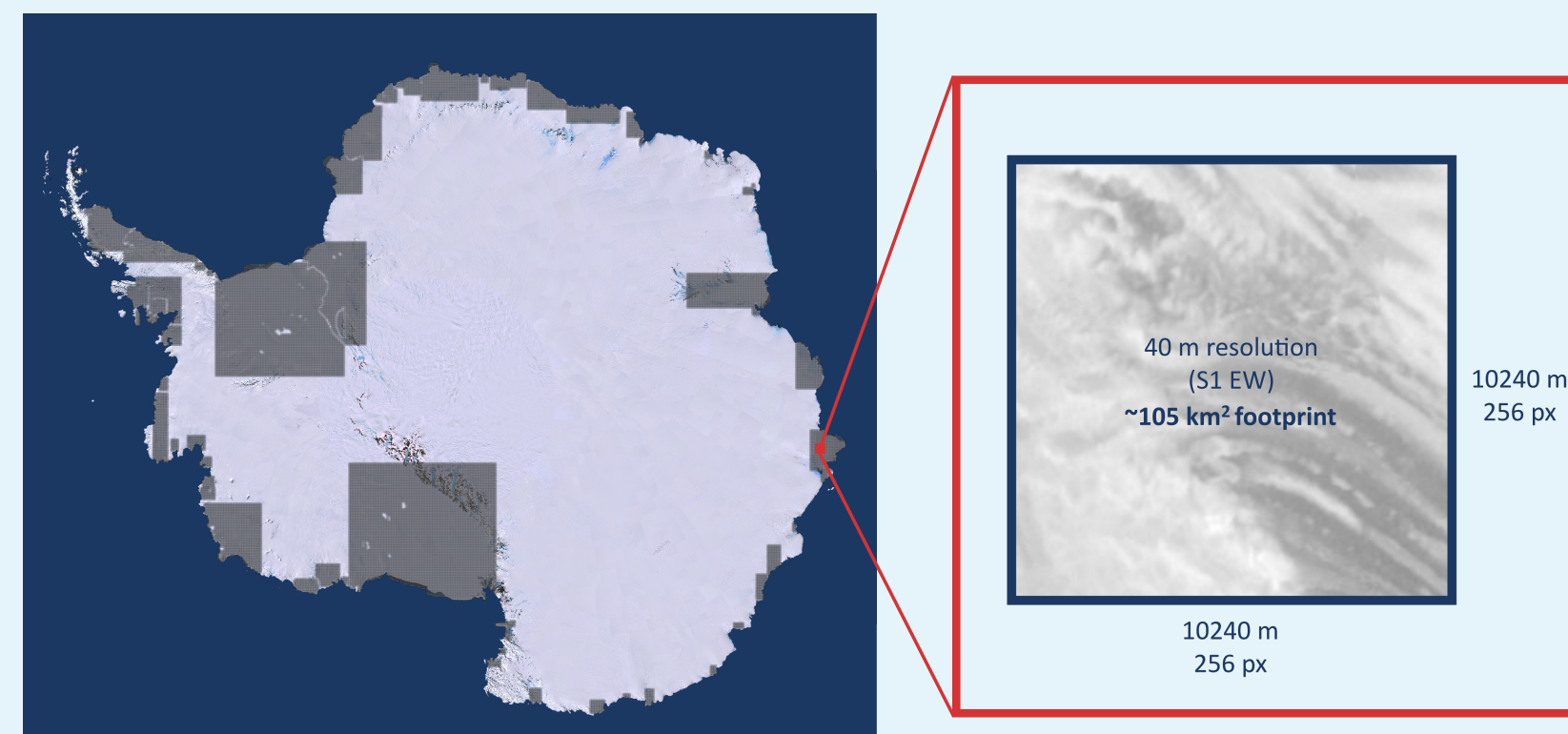
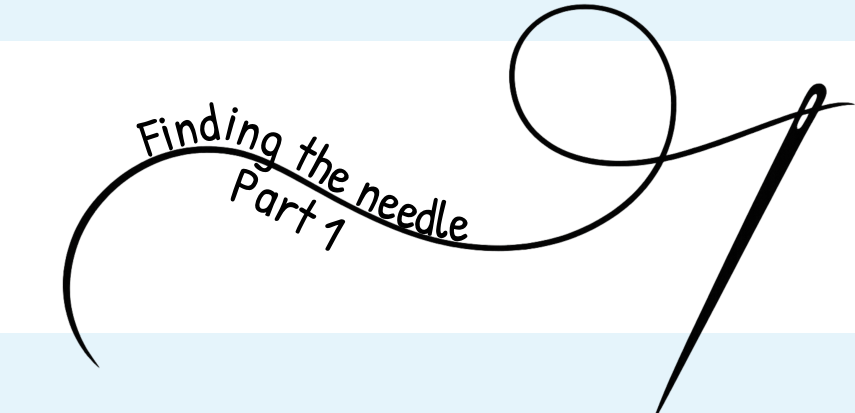


Figure 2: Sampling grids for Antarctic ice shelves overlain on LIMA mosaic (Bindschadler et al., 2008), zoom shows dimensions of one individual grid cell.

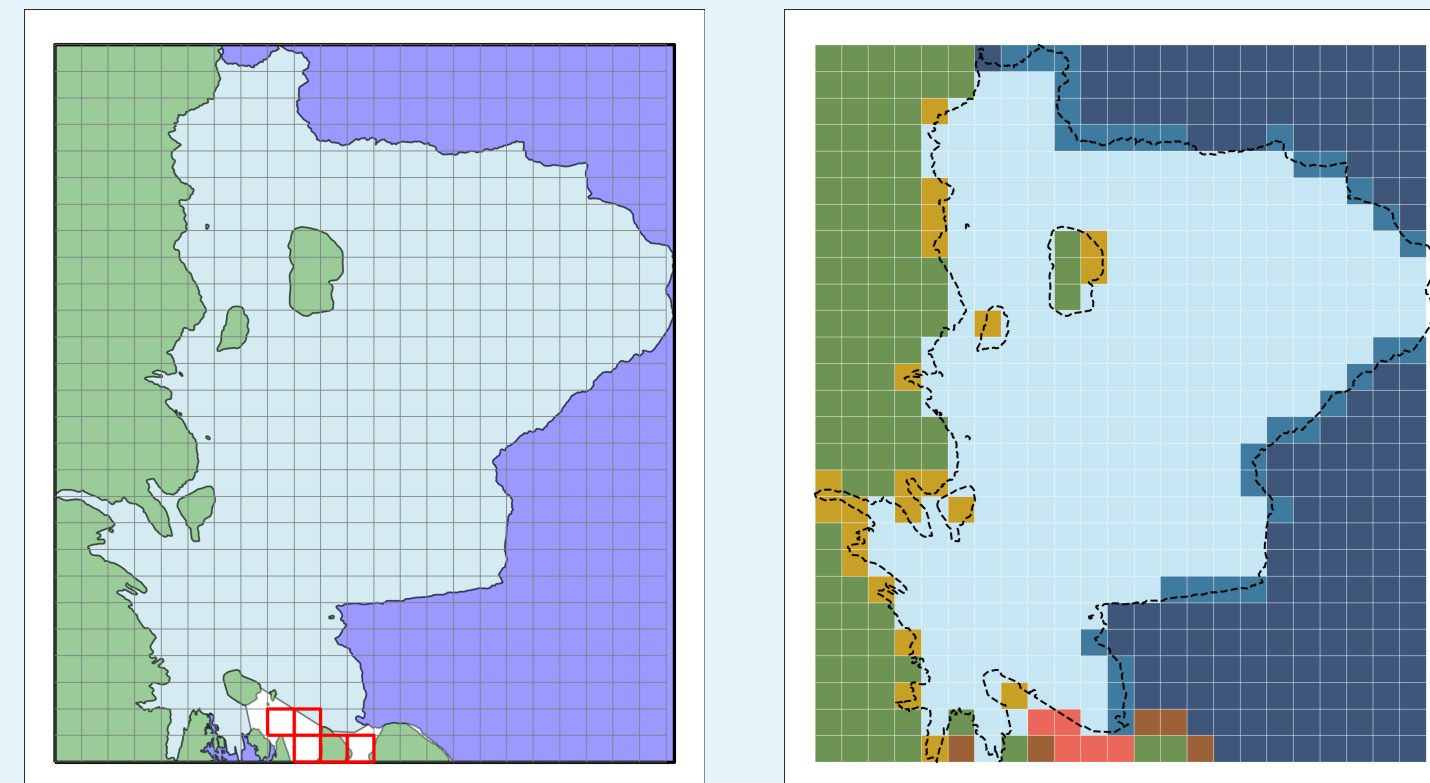


Figure 3: Exemplary grid classification for Shackleton Ice Shelf. Left shows land, ocean, and ice-shelf masks (Data from the SCAR Antarctic Digital Database, 2025), highlighted in red are grid cells belonging to neighbouring ice shelf. Right shows classified gridcells, where dark blue is ocean, mid blue is ocean margin, light blue is ice shelf, yellow is lateral margin, green is land, and brown is coastline.

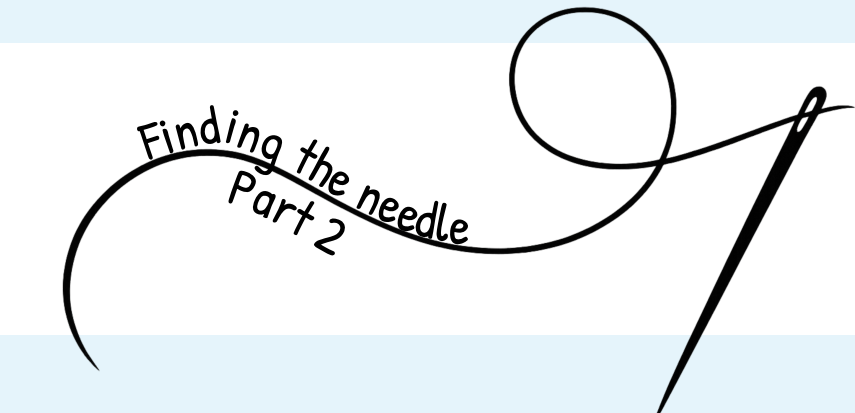
### Training data sampling:

- **Gridding** of ice shelf bounding boxes, disregarding ice shelves < 420 km<sup>2</sup> as well as ice shelves without Sentinel-1 SAR winter data availability
- **Classify grid cells** based on land, ocean, and ice-shelf mask, disregard open ocean and overlapping cells
- **Random sampling** 10% of available cells with forced equal weighting across 7 Antarctic regions
- **Enrich rare class** by manually sampling known buried lake locations (as well as sampling similar SAR signatures such as blue ice areas)

### Data preparation:

- **Sentinel-1 SAR** (HV polarisation) used to investigate meltwater ponding in the shallow subsurface [3][4]
- Two **seasonal composites** (early winter: Apr/May, late winter: Sep/Oct), mean composites to enhance data availability and minimise speckle noise
- **REMA DEM** added as topographic relief primary driver of meltwater distribution [2]
- Per-tile normalisation and **random year** assignment based on full seasonal coverage
- Labelling of SAR image tiles while consulting Sentinel-2 **optical image timeseries** to characterise melt regime

## Active Learning Framework and Model Architecture



- **Pool-based active learning** framework mitigates **labelling cost and scarcity**; the most informative instances to annotate are chosen from pool of unlabelled data
- **Human-in-the-loop** workflow: train model → class predictions on unlabelled data (inference) → select most uncertain instances to label → retrain model
- **ResNet-50 Encoder** with domain-relevant pretraining on Sentinel-1 SAR imagery (SSL4EO dataset) [5]
- **ASPP Bottleneck** gives multi-scale context, giving context to ambiguous patches of SAR imagery and allowing detected lake sizes to vary [6]
- **U-Net Decoder with skip connections** preserves high-resolution spatial detail, capturing even small compact lakes

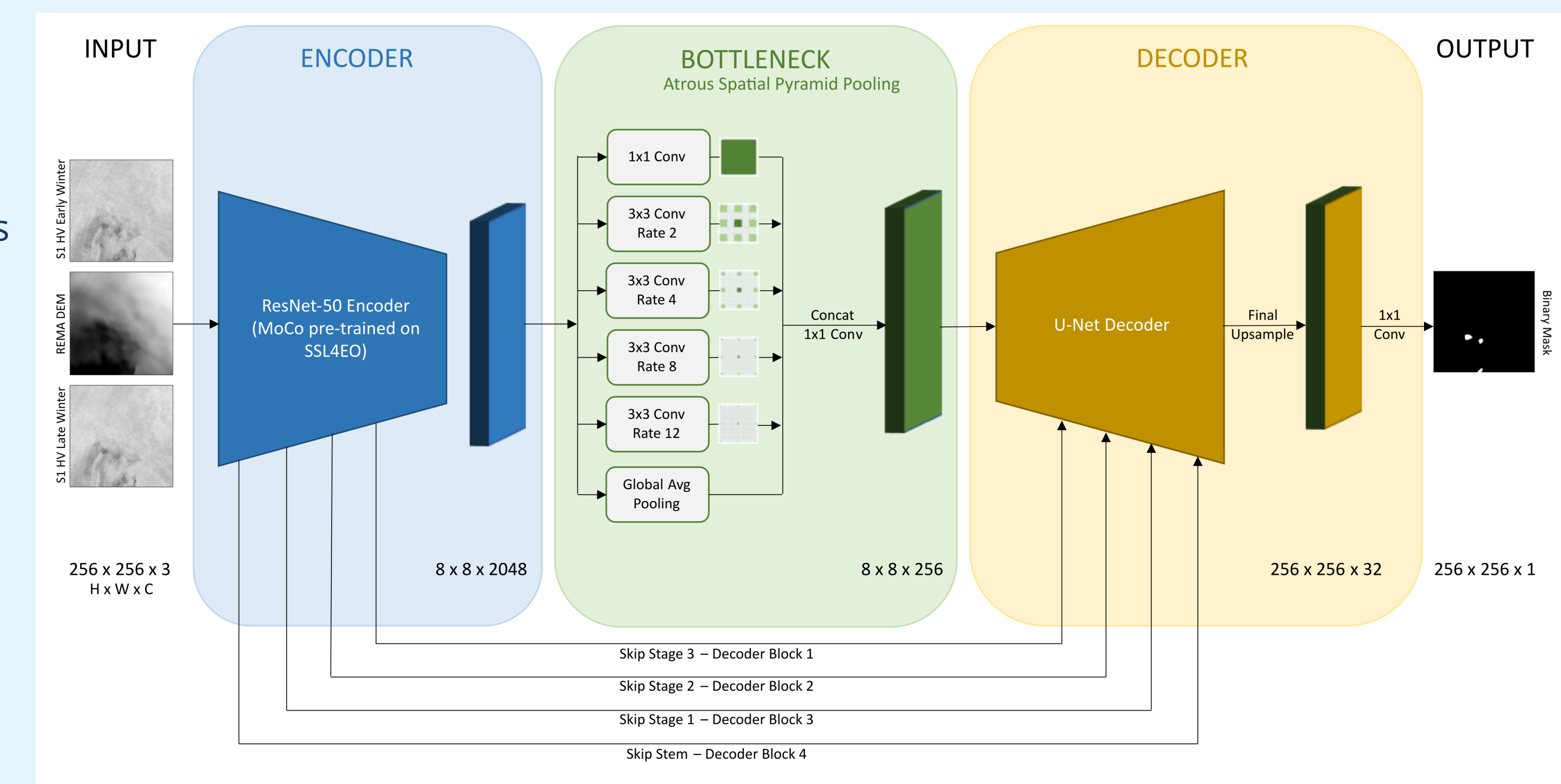


Figure 4: Proposed model architecture for BuriedLakeNet, a semantic segmentation model designed to detect buried lakes in Antarctica. Core architecture consists of in-domain (SAR) pre-trained ResNet-50 encoder, ASPP Bottleneck, and U-Net decoder. Model takes three channel input of two seasonal SAR composites and one REMA DEM tile, output is binary lake classification.