Detection of Melt Ponds on Arctic Sea Ice from Infrared Images using AutoSAM

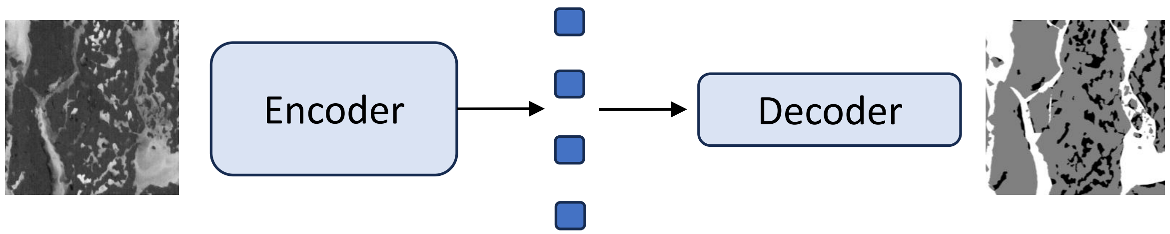
Marlena Reil[1,2], Gunnar Spreen[1], Marcus Huntemann[1], Lena Buth[3], Dennis Wilson[4]

[1]IUP, University of Bremen, [2]University of Osnabrück, [3]Alfred Wegener Institute, [4]ISAE-Supaero, University of Toulouse

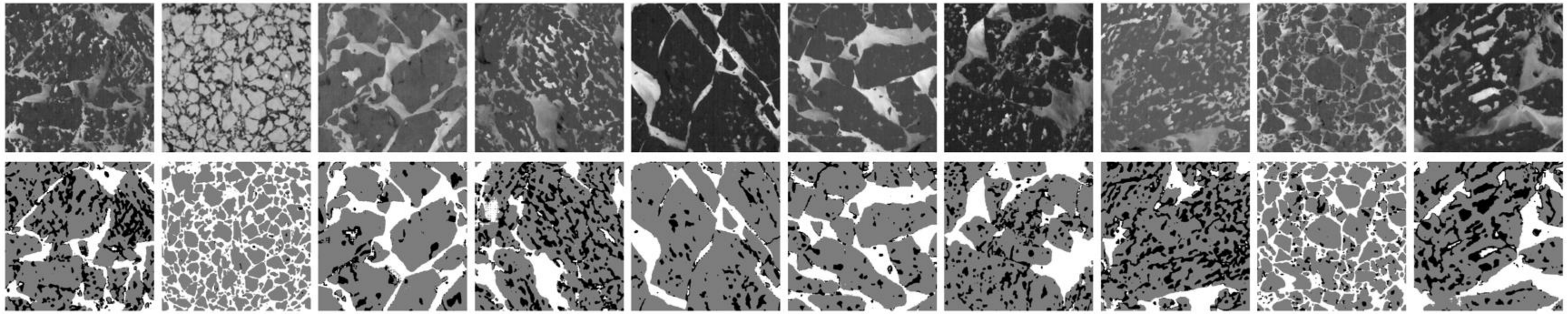
Motivation: Melt ponds are important components of the Arctic climate system but data are limited.

Goal: Extend existing melt pond data by segmenting helicopter-borne thermal infrared (TIR) images into different surface classes.

Method:



Results:

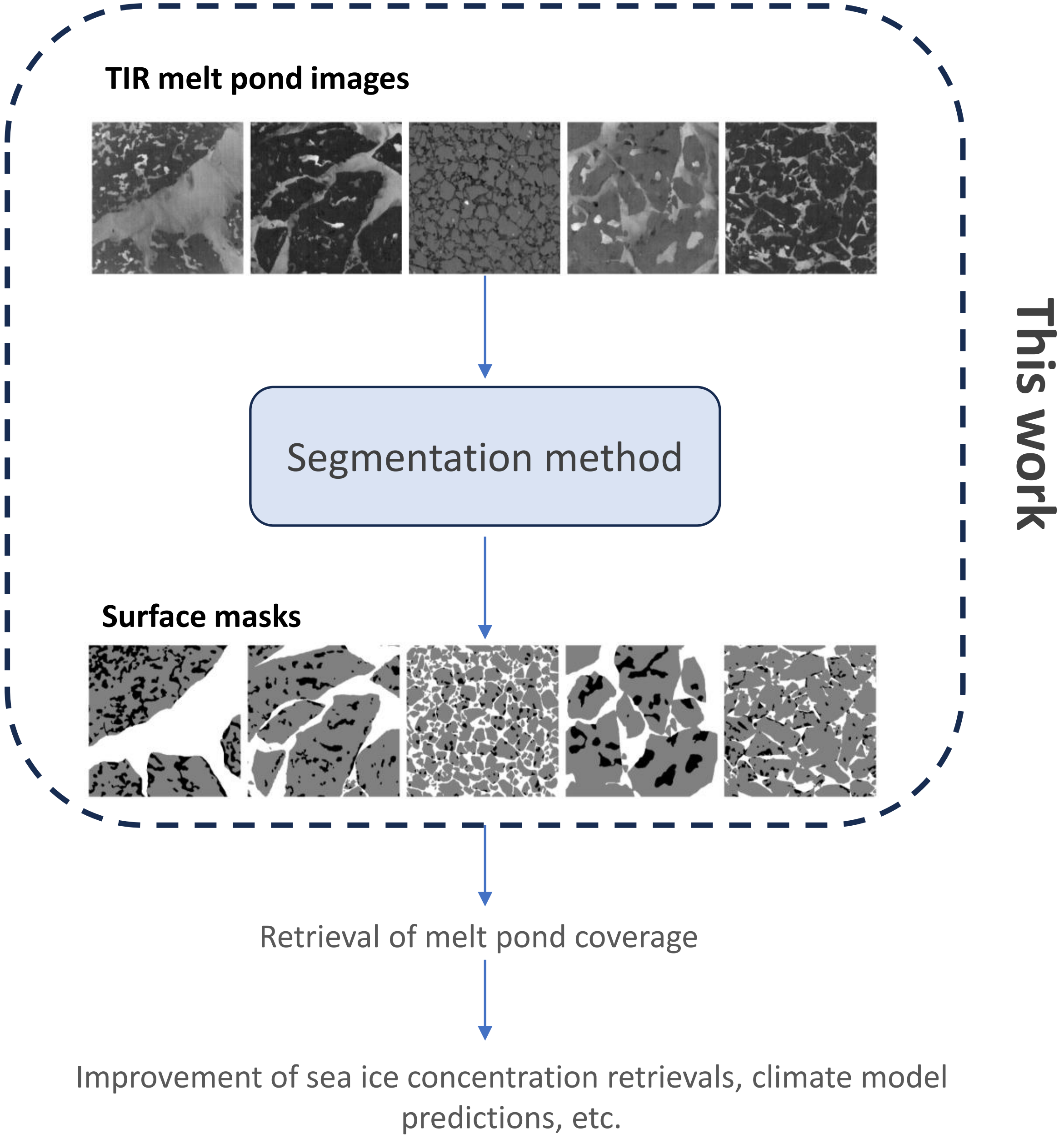


Melt pond
Sea ice
Ocean

Overview

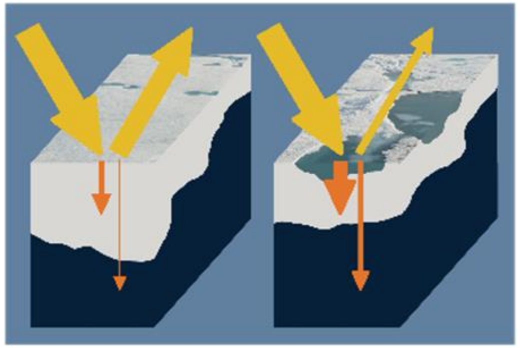
- Melt ponds are pools of water on Arctic summer sea ice that play an important role in the Arctic climate system.
- Retrieving their coverage is essential to better understand and predict the rapidly changing Arctic.
- Melt pond data are limited.

Here, we aim to **enhance melt pond data by segmenting high-resolution helicopter-borne thermal infrared (TIR) imagery into different surface classes.** We approach this using deep learning to handle temporally and spatially varying surface temperatures. Code, data, and models are provided online:

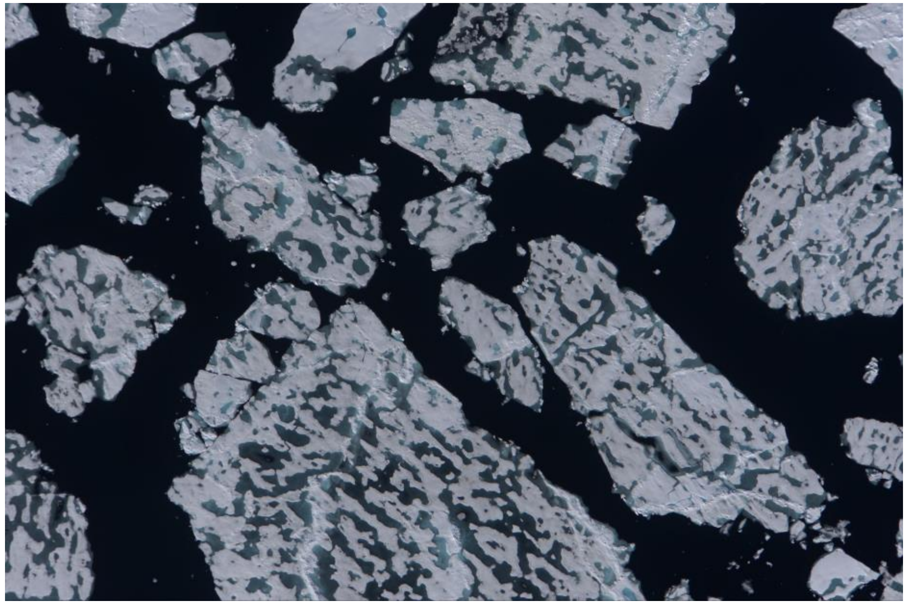


Melt Ponds

- Form on Arctic sea ice during summer as a result of ice and snow melt.
- Can cover up to 60% and 80% of the ice area.
- Range in size from square centimeters to square kilometers.
- Absorb significantly more sunlight than sea ice, causing further ice melt.
- Also impact the under-ice ecosystem by increasing light transmittance of the surface.
- Accurate measurements relevant to surface energy balance observations, improved sea ice concentration retrievals, and climate model predictions.



Melt ponds reduce the sun reflectance of the surface. Image: Hannah Niehaus.

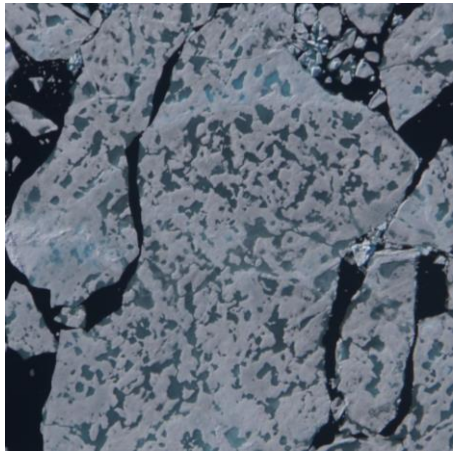


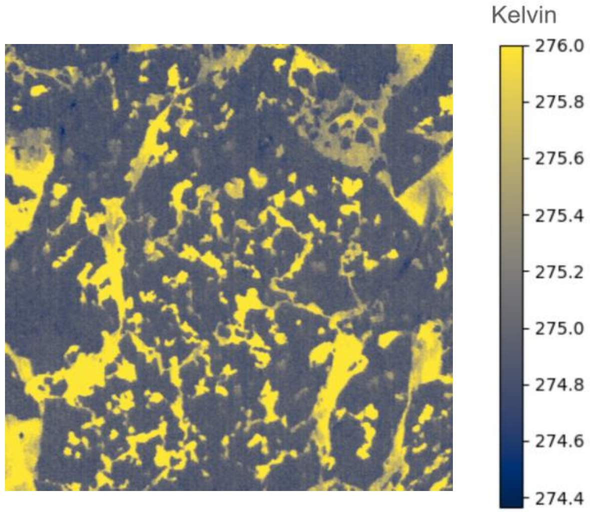
Helicopter image of melt ponds (AWI_PS131_02).

- Data are limited: In-situ measurements are rare and locally restricted, while most satellite products are too coarse to resolve individual ponds.
- Helicopter data as a compromise between scale and resolution.
- At the sensor level, existing work mostly uses optical imagery, which is dependent on daylight.

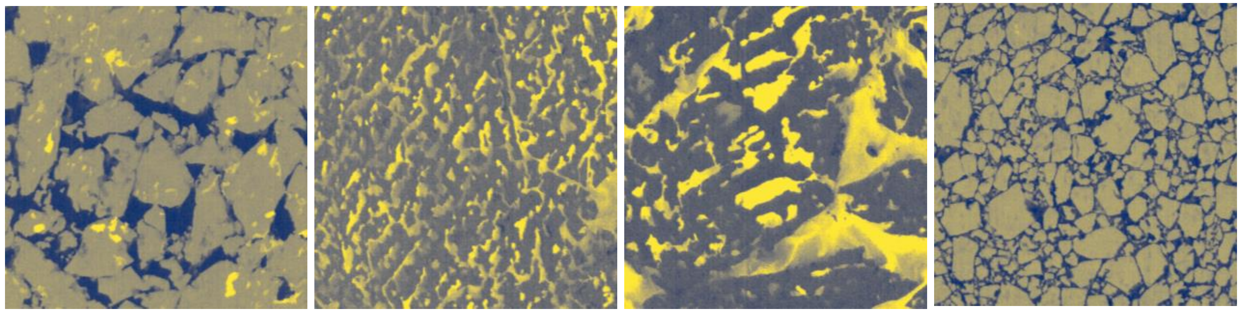
Data

- Broadband infrared radiation (7µm-14µm) at 1m resolution.
- 16 helicopter flights conducted in the marginal ice zone of the Fram Strait region in July and August 2022 (PS131 ATWAICE Campaign [1]).
- Gradient and drift corrected.
- 640 x 480 pixels per image.





Top right and bottom: Pseudocolored examples from our dataset. Optical image (top left) for reference (AWI_PS131_02). The temperature scale on the top right applies to all TIR images.



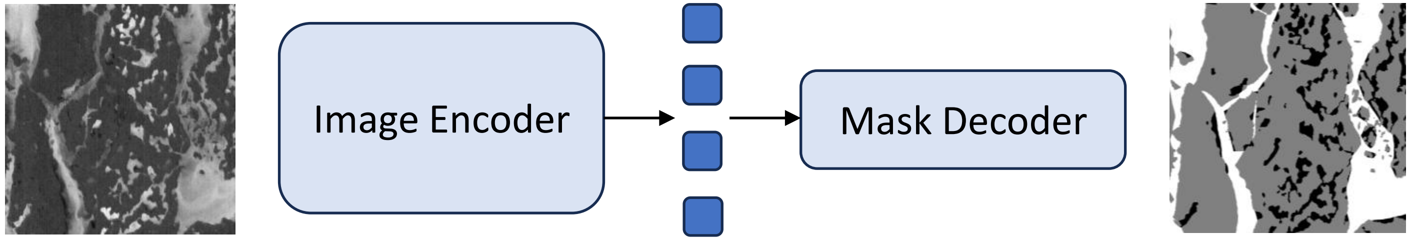
Advantage of TIR: Less dependent on daylight, allowing retrievals when the sun is low above the horizon or earlier in the season.

Challenge: Surface classes come in a variety of sizes, shapes, and temperatures. Traditional spectral or object-based segmentation methods are not applicable.

[1] Thorsten Kanzow. The Expedition PS131 of the Research Vessel POLARSTERN to the Fram Strait in 2022. Ed. by Horst Bornemann and Susan Amir Sawadkuhi. Bremerhaven, 2023. doi: 10.57738/BzPM_0770_2023.

AutoSAM

- Deep learning models are able to learn complex segmentation tasks by seeing large amounts of image/mask pairs.
- AutoSAM [2] is a deep learning architecture based on the segmentation foundation model Segment Anything (SAM) [3]. SAM is pre-trained on a large dataset of natural images (SA-1B).
- AutoSAM has been shown to perform well on medical images, which have similarities to our data domain.
- We hand-label 21 TIR images to adapt the pre-trained model to our dataset.



- The encoder transforms an image into contextual representations. We freeze the weights of AutoSAM during fine-tuning.
- The decoder generates an output mask from the encoded representation. We fine-tune the mask decoder on 11 of our manually annotated images and use the remaining 10 for validation.
- Both image encoder and mask decoder are based on Vision Transformer [4].

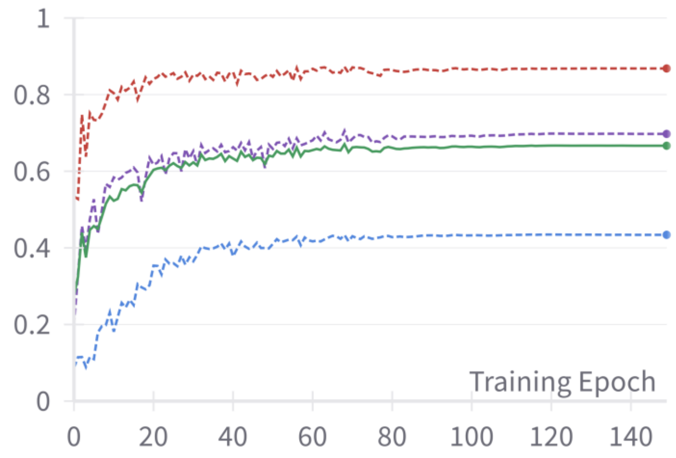
[2] SAM: Kirillov, Alexander, et al. "Segment anything." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.
[3] AutoSAM: Xinrong Hu, Xiaowei Xu, and Yiyu Shi. How to Efficiently Adapt Large Segmentation Model(SAM) to Medical Images. 2023. arXiv: 2306.13731 [cs.CV].
[4] Dosovitskiy, Alexey, et al. "An image is worth 16x16 words: Transformers for image recognition at scale." arXiv preprint arXiv:2010.11929 (2020).

Results

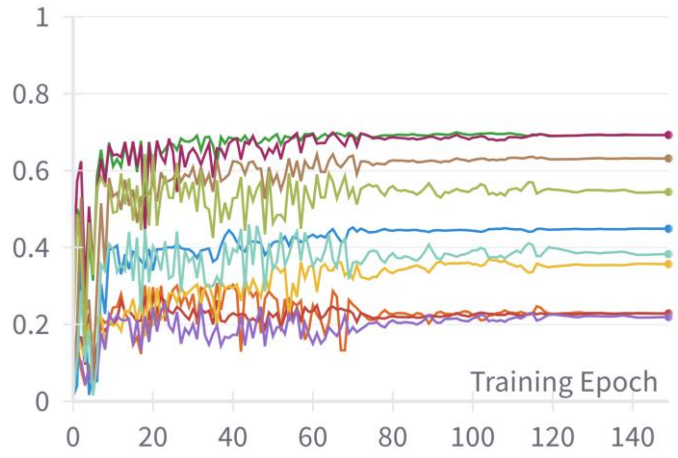
- We obtain an average Intersection over Union (IoU) of **0.667** and a Melt Pond IoU of **0.435**.
- AutoSAM outperforms U-Net and PSP-Net, two encoder-decoder architectures used for related tasks.
- High performance variability across classes: likely due to class imbalance in the training data.
- High performance variability across different validation images.

	Mean	Melt Pond	Sea Ice	Ocean
U-Net	0.582	0.320	0.823	0.602
PSP-Net	0.499	0.230	0.779	0.488
AutoSAM	**0.667**	**0.435**	**0.868**	**0.698**

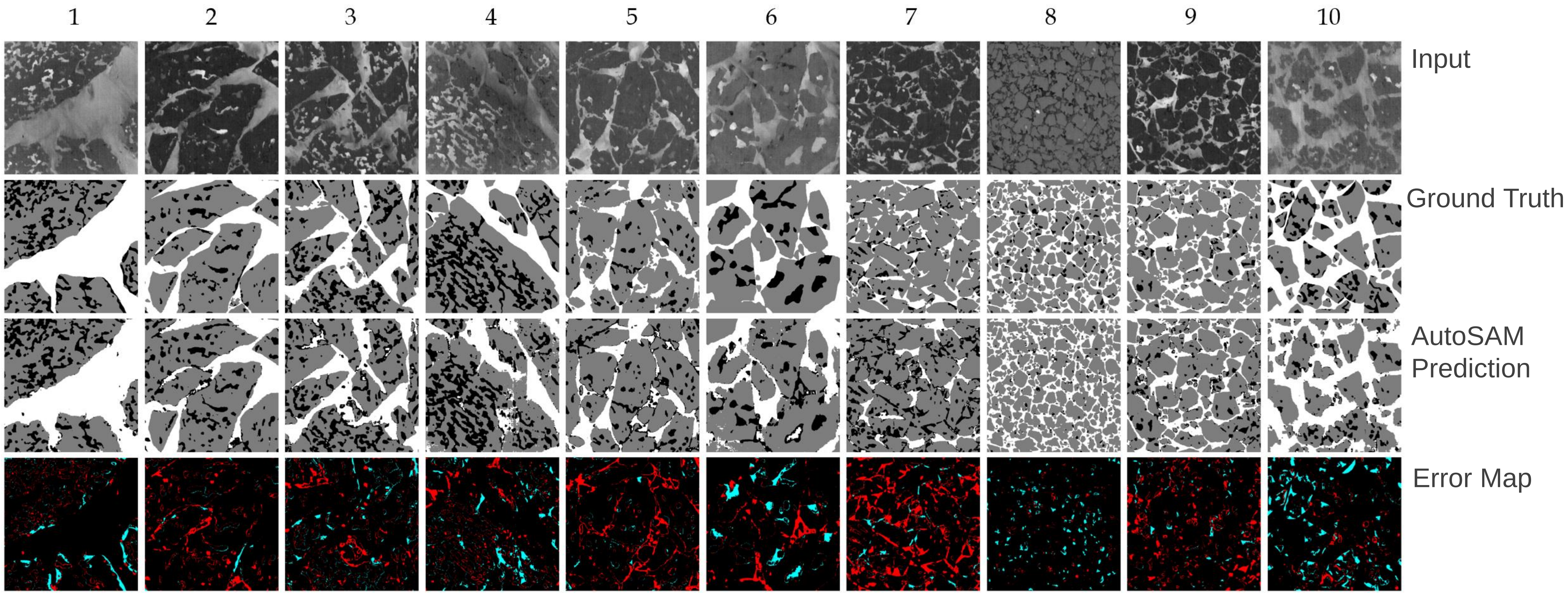
Mean and per-class performance measured in IoU.



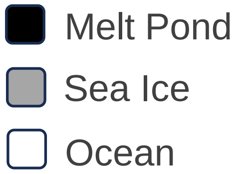
Per-class IoU for different fine-tuning epochs. Red = Sea Ice, purple = Ocean, green = average, blue = Melt Pond.



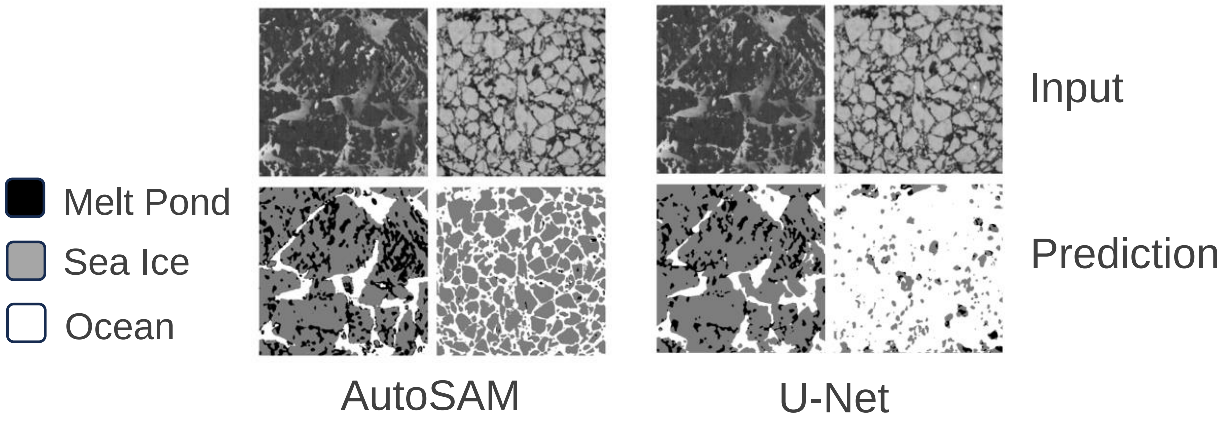
Per-image performance measured in Melt Pond IoU. Different colors refer to different validation images.

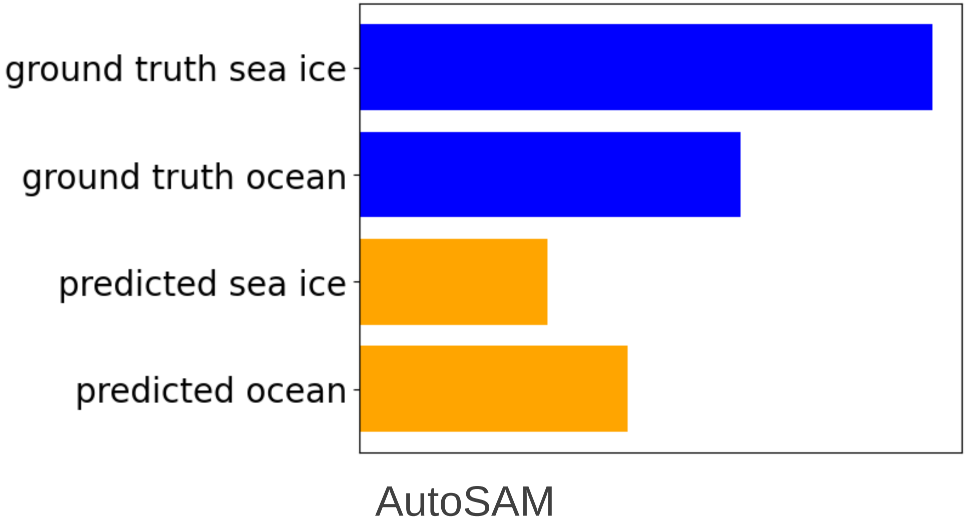


Error map: Shows melt pond false positives in red and melt pond false negatives in blue.



- Good performance on large ice floes with visually well separated surface classes.
- Weaknesses: melt ponds on smaller floes, misidentification of ocean gaps between floes (e.g., Samples 6, 7), fuzzy boundaries (Sample 10).
- Incorrect prediction of ocean within correctly delineated melt ponds (e.g., Samples 4, 6).
- Overestimation of melt pond boundaries (see error map).





- Compared to U-Net, AutoSAM can capture surface types independent of their relative temperature differences.
- Melt pond false positives (blue) are higher than melt pond false negatives (orange), which suggests a general overestimation of melt ponds.

Conclusion

Code and data *Bachelor thesis* *OSPP rating*

Summary:

- We addressed the lack of melt pond observational data by hand-labeling helicopter-borne TIR images and fine-tuning AutoSAM for surface class segmentation.

Achievements:

- AutoSAM can predict varying surface conditions in TIR data.
- Further development of the method contributes to the incorporation of TIR into melt pond analysis, further enabling light- and season-independent study of the Arctic surface.

Limitations:

- Results are preliminary. Model performance scales with size of training data, but manual annotation is time-intensive.
- We excluded images with poor visibility for annotation and prediction, so our results may not be representative of the entire dataset.

Future Directions:

- Simultaneous optical images are available. VIS/TIR fusion would allow expansion of the training data and multimodal parameter retrieval.

Contact: marlena1@gmx.de.