**IMPLEMENTATIONS:**

* **K-crossfold**
* **Pretraining autoencoder**
* **Hyperparameter optimization**
* **Vis images**
* **augmentation**

**IN THE END:**

* **Check related work**
* **Irgendwo semantic segmentation introducen**
* **CNN und semantic segmentation part**

**Heute:**

* **Original extraction**
* **Augmentation (verify augmentation steps)**
* **Pretraining implementen**
* **K-cross validation**

**BREMEN 01.06.**

<https://www.sciencedirect.com/science/article/abs/pii/S0165232X15001433> access

show results of two images with new plotting method

is TIR only for lower latitudes still true?

**QUESTIONS:  
- Kann ich Github Repos citen**

* **An welcher Stelle Zusammenfassung was ich mache (Related Work?)**
* **Supervised / unsupervised erklären?**

**TO-DO:**

* **Implement pretraining:**
  + **Idea: Build an autoencoder on IR images and use weights for pretraining UNet. Investigate on 1-channel 🡪 3-channel conversion**
  + **Run autoencoder for 1000 instead of 500 epochs 🡪 better results.**
  + **Change last covn2d layer of autoencoder to 1 filter to yield grayscale image result**
  + **Autoencoder will be build the same way as UNet (but with skip connenctions in decoder)**
  + **Pretraining autoencoder saves time but doesn’t necessarily improve performance???**
  + **The final training images are not part of the pretraining dataset**
  + **Improvement for later: augment, batch,… also pretraining**
  + **Mean squared error as loss, and sigmoid as activation**
  + **Freeze encoder because images are similar (same features)**
* **Implement cross-fold validation**
* **Implement hyperparameter optimization**
* **Augmentation methods**
* **Check UNet implementation, reproduce results**

**Issues:**

* **How to handle 256 / 480 prediction**
* **Check normalization / batch normalization in beginning**

**Minutes Ulf 02.06.2023:**

* **IMPORTANT: normalization on whole image / set (not single patches)**
* **Normalization muss mit model / aktivierungsfunktion zusammenpassen**
* **Pretrain with different models and discuss**
* **Results: predict with smaller patch size und take majority (auch um varianz 32 zu checken)**
* **Augmentierung: Skalieren, zoom macht vielleicht doch sinn**

**Meeting Ulf:**

* **Backbone for pretraining**
* **ResNet34 Backbone: Implementierung mit Batch\_Norm am anfang 🡪 kein preprocessing im Vergleich zu keras ResNet50 Implementierung. Habe Angst, dass model nicht läuft wenn ich andere Implementierung nutze.**
* **Warum überhaupt pretraining auf gelabelten Datenset?**
* **Welches ImageNet**
* **Welches RS pretraining dataset (fühlt sich ein bisschen random an), BigEarthNet? SkinLesion?**
* **Brightness Change und crop als augmentation methoden?**
* **Cross validation und hyperparameter optimization schwierig**

**Meeting 26.05. Minutes:**

* **Cross validation for one hyperparameter set**
* **Use their dataset and reproduce results**
* **Autoencoder**

**Meeting Ulf 23.05. Minutes:**

* **Teil zu Schmelztümpel, Remote Sensing: Warum interessant, physikalische Aspekte in Bezug auf IR, warum kann man sie in IR gut sehen, typische Formen (falls relevant für Arbeit, optisch kleinere zerfranzte Strukturen und Formen, Aspekte die interessant sind für die Bildverarbeitung) 🡪 in Einleitung oder auch Forschungsstand (da und da gibt’s Probleme…)**
* **Related Work (Theorieteil / Forschungslandschaft): Methoden erwähnen und wo stehen wir.**

**Was ist schon gelöst und was noch nicht und welche Ergebnisse kann ich erwarten, was ist noch unklar (auch Bezug Eissegmentierung)**

**Warum Unet für diese Aufgabe, Stärken Schwächen und mögliche Konsequenzen für das Ergebnis.**

**Welche Augmentierungen sind sinnvoll / nicht sinnvoll**

**KLEINER DATENSATZ und Methoden um da heranzugehen: Transfer Learning, was hat verschiedenes Domain für nen Einfluss, patches aus homogenen Datensatz…**

**MEINE Details abarbeiten, welche Problemen und Chancen. UNet nur zusammenfassend, sonst eher auf relevante Quellen verweisen (nicht mein Hauptthema).**

* **Methoden: Methoden konkretisieren, beziehen auf Theorieteil möglich**
* **Uni Grid: Git Bash, auf remote conda environment installieren (Ulf hat Slides)**
* **Coding:** 
  + **Mit 60 epochen könnte man beim training aufhören, wobei große models manchmal noch bisschen länger trainieren…**
  + **Test acc höher als train acc 🡪 sind Dropout und augmentierung (train) dran schuld?**
  + **Cross-validation statt splitting nehmen**
  + **Evaluierungsmethoden: manuell macht schon auch sinn**
  + **128 als patch size**
  + **Uneindeutiges ergebnis masken 🡪 was genau ist der output? (winner takes it all macht vllt sinn oder auch gewichtete Summe)**
  + **Transfer Learning: Wär schon cool, auf welchem Imagenet ist vortrainiert (segmentierung oder klassifizierung? Nur encoder oder auch decoder?). How: Herunterladen, trainieren, freezen.**
  + **AUGMENTIERUNG GENAUER: welche machen sinn und welche nicht so, rotieren anschauen, scalen, …**
  + **Optische Daten Ergebnisse: Warum ist der Ground Truth falsch?**
  + **Evntl. Matching optimieren**
  + **Check exact architecture unet, sm unet**

**Meeting Ulf Notes:**

* **Kann ich nach 60 epochen aufhören mit trainieren?**
* **Evaluierungsmethoden (iou, loss, manuell?)**
* **Transfer Learning: wie funktioniert das? (muss ich datenset herunterladen, dann trainieren und layers freezen)**
* **Test accuracy höher als train accuracy**
* **Eher cross validation als train-test-split, wegen kleinem datensatz**
* **Warum sind die ergebnisse nicht eindeutig? (printed er die wahrscheinlichkeiten?)**
* **Aus Meeting mit Gunnar: 128 als batch size**

Plan:

Heute: Introduction Melt Ponds und Remote Sensing

Morgen: Segmentation Models adjustment + hyperparameter selection, Background study

**Meeting Gunnar 22.05.:**

**Evaluation: Zeitvergleich VIS (kmeans) und IR, MP Fraction entlang des Fluges ist Ziel. Auch aus anderen Flügen Bilder testen.**

* **Nochmal schauen: was predicted er da? (keine klaren Klassen), interpoliertes Ergebnis?**
* **128 als Patch Größe (kleinerer Patch, mehr daten)**
* **Pretraining (Coxi interessant)**

**Zeit Automatisch:**

* **Kmeans und mergen: 16937325.15668869 ms**
* **Matlab: 16 Bilder in 1,5 Stunden**
* **Lorenzo Process: 2m30, 3m30, 1m46, 1m46, 2m3,** 171186.55800819397, 126069.15950775146, 148977.22959518433, 204942.21949577332, 163412.5895500183, 126553.09581756592, 214785.89534759521, 226176.98001861572, 179828.5892009735, 165841.78638458252, 174129.28175926208, 174752.90417671204, 223262.4592781067, 169996.88625335693, 173352.79774665833, 167928.14135551453, 189038.7740135193, 171975.77166557312, 157485.76283454895, 125603.3308506012, 138408.9958667755, 122463.78755569458, 120686.71464920044, 125419.92378234863, 125643.1827545166, 132097.47433662415, 133841.85600280762, 131505.52129745483, 149458.1422805786, 149909.94095802307, 133986.18745803833, 129329.56838607788, 125629.16851043701 (nach 39)
* [144251.27696990967, 126088.79041671753, 148644.42086219788] (47,48,49)

**Meeting Montag:**

* Hyperparameter Optimization (epochs, loss, batch\_size, augmentation…), decide on metrics & Co
* What exactly should be the outcome (calculate melt pond fraction?)
* Evaluation: error rate?

Open Questions:

* How big are the scenes?

**Dann**

* Think about train setting (hyperparams, augmentation)
* Decide for VIS images (idea: take 10 of each vis, hsv and gray and show differences)
* Computing grid IKW

**Hyperparameter Grid Search**

Epochs, backbone, batch size, learning rate, optimizer

**Melt Pond, sea ice and ocean properties from thermal infrared imaging during Arctic summer**

* “ice surface properties are subject to strong and rapid changes.”
* “surface albedo gets reduced, which has strong impact on the surface energy balance, and ocean – ice – atmosphere interactions are increasing”
* Thermal infrared = IR
* ‘to analyze the variability and evolution of surface properties like melt pond and sea ice fraction, flow size distribution, or ocean and ice surface temperature’
* were taken from helicopters (good spatial coverage) and from onboard the ship (good temporal coverage) and will be used to better understand the summer sea ice and ocean interactions and later improve satellite retrievals
* the different surface types need to be automatically separated. As the surface temperatures of the different types change spatially and temporally, simple thresholding might not be sufficient and ML approaches should be a suitable solution.

<https://seaice.uni-bremen.de/melt-ponds/>

* MP affect the albedo, mass balance and heat balance of ice by translating air temperature increase into drastic and rapid surface type changes.
* Introduce a positive feedback within the sea ice albedo feedback loop, facilitating further ice melt.
* Knowledge of melt pond fraction and spatial distribution is therefore of major importance in the context of changing climate.
* MP fraction

Abstract

Melt Ponds are crucial to understand the heat budget of the Arctic region.

‘The objective of this work is to analyze and perform dense pixelwise segmentation on the images to automatically compute the melt pond fraction’.

This work investigates the automated segmentation of melt ponds from Helicopter IR images during Arctic summer. Experiments show that the resulting method could achieve an IoU of 90%.

* MPs cover up to 50-60% of the sea ice area (<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2011JC007869>)

Introduction

The Arctic region warms three times faster than the rest of the world. Melting snow and sea ice cause the formation of melt ponds on Arctic sea ice. Melt Ponds are pools of water that form on top of sea ice when temperatures are rising close to zero degrees. As temperatures are rising close to zero degrees in summer and sea ice is melting, pools of water form on top of the ice – surface structures known as Melt Ponds. Melt Ponds have a significant impact on the reflectivity of the surface. As they consist of darker-colored water, they absorb more sunlight compared to brighter sea ice. This leads to further warming and accelerated melt of surrounding areas, resulting in a positive feedback loop. An improved understanding of the concetration of MP and development over time would… We require accurate estimates of the surface concentrations. Providing a robust method to analyse melt pond fraction which allows to study melt pond concentration and variation over time is therefore of major importance to understand the Arctic heat balance and potential consequences for global climate change.

Due to the remoteness of the area, melt pond data is limited and restricted to satellite imagery and airborne images. Satellite data can provide a temporal resolution of a few days but lack of geometric accuracy. On-side data exists but lacks of spatial coverage and is challenged by further geometric distortion. Airborne images provide a good trade-off of geometric resolution but data is limited to a few flights per expedition. Using airborne images to derive melt pond fraction can provide accurate results and later serve as ground truth to improve satellite retrievals. ‘computation of temporal and spatial ice distributions can help to validate satellite retrievals’.

‘Current methods of estimating pond coverage on sea ice from MODIS have high error and bias when compared to new high-resolution datasets’

While Melt Ponds are often studied with optical bands, IR data exists.

To retrieve the fraction of melt ponds of image data, different surface types need to be extracted.

To extract image data into different surface types, image segmentation methods are reasonable. While optical data can be extracted based on surface color, IR images are more difficult to segment.

Temperature

Temperatures vary across the surface types and shapes must be taken into account. More complex methods that can extract features must be taken into account.

Background Study and Related Work

Deep Learning for Semantic Segmentation has been studied a lot within the last decades. With increasing importance of feature extraction methods as Convolutional Neural Networks,… Common architectures

Segmentation applied to studying melt ponds is still limited in literature. … constructed a U-Net to study Melt Pond fraction.

Deep Learning Segmentation is applied in Remote Sensing tasks. Architectures that are often studied are…, …, … .

This work is concerned with

As temperatures rise close to zero degrees in summer, sea ice melts and melting water collects on areas of lower surface elevation, forming melt ponds. Melt Ponds significantly change the reflectance behaviour of the surface and accelerate further melting of surrounding areas.

Understanding the fraction and development of melt ponds is therefore of major importance.

The melting of Arctic sea ice also contributes to the amplification of global climate change. Ice has a high albedo, meaning it reflects a significant portion of incoming solar radiation back into space. As the sea ice retreats, darker surfaces such as open water and land absorb more heat, leading to further warming and accelerated ice melt—a phenomenon known as the ice-albedo feedback. This positive feedback loop intensifies the rate of climate change not only in the Arctic but also globally.

Due to the remoteness of the area, studying melt ponds is often restricted to satellite imagery and limited on-side and airborne data. Whereas satellite images can provide a high temporal resolution

Studying melt ponds is challenged by remoteness of the area. Common techniques are Remote Sensing, collecting data through satellites from space, airborne vehicles or on board measurements. While satellites are able to provide a high and regular temporal resolution and constant parameter adjustments, helicopter images are able to provide a better geometric resolution and are especially suited for extracting objects with the size of melt ponds.

While common wavelengths measured are optical bands, Radar, also IR data exists. Mapping the temperature values allows a more intense study of the measured area and predictions about future surface temperature. Temperature anomalies might allow to predict melt ponds of the upcoming season.

To derive melt pond fraction from IR images, image segmentation is crucial

Deriving melt pond fraction from remote sensing images is challenging due to temperature inconsistencies. Salt water might be colder or warmer than surrounding ice, melt pond. Classification not only based on color but also melt pond shape is crucial.

While in optical images, different colors allow for simple, pixel-wise color segmentation, IR is more difficult. Machine Learning methods might be able to tackle this.

Introduction

* Why melt ponds
* Why helicopter images of melt ponds
* Why IR
* What methods are there to extract
* Why Deep Learning

Theoretical Background

* Melt Ponds
  + Definition of Melt Ponds: Appear in summer months
  + Effects on climate system
  + How melt ponds are studied
* Remote Sensing and IR
* Image Segmentation and related work
* Data Annotation
* UNet

Discussion

* Melt Ponds vs leads
* More diverse training data
* Images need to be undistorted

**Melt Ponds**

Melt Ponds are significant features that form on the surface of Arctic Sea Ice during the summer months. They occur as a result of melting of snow and ice, creating pools of water on top of the ice. Melt Ponds play a crucial role in the energy balance of the Arctic and affect the overall ice mass and interaction with the surrounding environment. Understanding melt pond distribution is of importance to understanding the Arctic climate change and potential consequences.

Melt Ponds have a significant impact on the reflectivity of the ice, as darker-colored water absorbs more sunlight compared to reflective ice. The increased absorption of solar radiation leads to further melting and creates a positive feedback loop, leading to melt pond sizes of … by the end of the melting season.

Studying Melt Ponds includes in-situ measurements of depth, evolution and distribution over the summer months. Moreover, Remote Sensing techniques are applied including satellite imagery, airborne sensors. While satellites can provide a higher temporal resolution and more stable parameters, airborne techniques allows for a more finegrained geometrical investigation.

Unlike VIS or RADAR imagery, that allows for structural analysis, Infrared wavelength allows to study the temperature distribution of melt ponds. Comparing winter and summer measurements allows for prediction of melt ponds from winter images and might lead to better predictability (Linda). Classifying melt ponds from IR images is therefore of major importance.

Extracting melt ponds from infrared images provides an additional challenge. In the infrared domain, surface temperatures are changing. Melt Ponds can have the same temperature as surrounding ocean, colder than ice (temperature of melt ponds).

Extracting melt ponds from images by temperature only is therefore not possible and simple thresholding methods (like used in … for example) might not help. The shape needs to be taken into account.

<https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JC02034>

<https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2000JC000583> (MP evolution)

**Segmentation**

Artificial Neural Networks are widely used, also in Remote Sensing. Studying Melt Pond fraction so far has been come down to the following techniques… Artificial Neural Networks have been used in works like…, …, … . However, to the best of the authors knowledge, there is no technique yet that can use IR images as sufficiently input.

Segmenting Melt Ponds from Infrared images is an issue to being regarded so far, to the best of the authors knowledge. Common segmentation methods, separating images into different semantic classes is a common topic in Remote Sensing and Sea Ice Research. Methods often include setting brightness thresholds, color-based segmentation or kmeans methods. In … et al. used this approach, in … et al. used that approach.

In the last decades, Deep Learning methods have been proved successful for Segmentation Tasks. Convolutional Neural Networks (CNN), there are single approaches that use LSTM or GANs. While ANNs are able to automatically extract features.

As Neural Networks are supervised techniques, they require large amounts if labelled training data. As the methods learn from the data provided, labelling quality, sufficiently large dataset, and diversity is of major importance. Manual labelling might yield more costumizability but takes a lot of time and resources. Automated approaches might help to fasten the process. As IR images are hard to use themselves as basis, corresponding VIS images are taken into account. In VIS images, surface types can be extracted more easily by color, automated unsupervised segmentation methods (kmeans) can be used. This allows to create a larger dataset, but might be prone to small errors. Additionally, distortions between the image types need to be taken into account.

This work trains and evaluates three different Convolutional Neural Networks (UNet, PSPNet, DeepLabV3+) that extracts IR Helicopter Images into Ocean, Sea Ice and Melt Ponds. Two different labelling methods are tested: Small, accurate, manually labelled and larger, automatically labelled from corresponding VIS images. The goal of this work is to create an accurate training dataset for model. As an ablation it is tested if a larger, automatically extracted VIS dataset is able to outperform a small, manually created dataset, enlargened only by augmentation methods.

The hope is, that this work can be extended to different flights, also on-ship data. Future work could extend this to winter dataset, too.

The reminder of this work is divided in the following parts: In section … I will present the methodology used, methodology explained in more depth. In experiments the experimental setting and results. In Discussion I will critically evaluate the outcomes and provide a summary of the findings in Conclusion.

Extracting melt ponds

<https://tc.copernicus.org/articles/6/431/2012/>

<https://epic.awi.de/id/eprint/38709/1/1-s20-S003442571500108X-main.pdf>

<https://tc.copernicus.org/articles/9/1551/2015/> (algorithm comparison ship, heli, satellite)

<https://www.cambridge.org/core/journals/annals-of-glaciology/article/melt-pond-distribution-and-geometry-in-high-arctic-sea-ice-derived-from-aerial-investigations/8770BFD400443CB8704AE5B32D3577AE>

<https://www.cambridge.org/core/journals/annals-of-glaciology/article/comparison-of-different-retrieval-techniques-for-melt-ponds-on-arctic-sea-ice-from-landsat-and-modis-satellite-data/3E7FBFA42BDBD8EDD84778C13FF44BE3> (MP fraction with PCA)

**Related Work**

Melt Pond Fraction is…

* Applications
* Methods for extraction

<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2015JC011030> (auch für Intro)

<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2011JC007869>

Wo stehen wir in der Melt Pond Fraction Zeitreihenanalyse aus Satelliten und optischen Daten. Warum wäre MP fraction wichtig. To accurately understand MPs, also thickness… would be necessary. This work only considers the task of obtaining estimates of surface concentration of melt ponds from digital image data. (manual analysis impractical).

<https://www.tandfonline.com/doi/abs/10.1080/00040851.1997.12003254>

<https://www.nature.com/articles/nclimate2203>

<https://www.sciencedirect.com/science/article/pii/S003442571500108X?casa_token=eRW0VRkFj1kAAAAA:U26nUONEDcdacIgj7enkwZXe1M3RprVzYJNp8bEAFAhMpH7PbdGd2Z8bI4PdE1fBqKqFSx_CNXwu>

https://tc.copernicus.org/articles/6/431/2012/

<https://tc.copernicus.org/articles/9/1551/2015/>

<https://ieeexplore.ieee.org/abstract/document/6784072>

**Algorithms for MP retrieval**

[Sea Ice Melt Pond Fraction Derived From Sentinel‐2 Data: Along the MOSAiC Drift and Arctic‐Wide - Niehaus - 2023 - Geophysical Research Letters - Wiley Online Library](https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GL102102)

[A New Algorithm for Sea Ice Melt Pond Fraction Estimation From High‐Resolution Optical Satellite Imagery - Wang - 2020 - Journal of Geophysical Research: Oceans - Wiley Online Library](https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019JC015716)

Melt Ponds are special due to their shape (related work there). They can build smaller channels between and melt completely through. Their size can reach… extending the whole image scale. Physical properties. Why visible in IR and use for research (Lindas Work).

Pond geometry

<https://ieeexplore.ieee.org/abstract/document/6784072> :

‘Melt ponds are typically interconnected on first- and second-year sea ice but discrete on multiyear ice or heavily deformed ice due to topography’ (3)

‘The size of melt ponds ranges over two orders of magnitude; small melt ponds with a size below 50 m2 are found to be most common‘ (4)

‘Melt Ponds heavily influence the accuracy of microwave radiometer ice concentration retrievals / inclusion of melt ponds in models has been shown to provide more realistic progress of summer melt’ (5,6,7)

Segmentation of Melt Ponds of Image Data from IR has been unexplored so far, to the best of the authors knowledge. Segmentation has been done from optical images.

What is different and challenging from extracting from IR is colour distribution. 🡪 DL 🡪 lack of training data (very time-consuming and extremely difficult, even for experts and with ground truth data available).

(DL has been used for related tasks such as ice segmentation or land cover / medical analysis)

Segmentation of similar tasks (only sea ice – ocean) and with which methods

River DL: <https://ieeexplore.ieee.org/document/9063496>

Semantic Segmentation DL MPs: <https://ieeexplore.ieee.org/abstract/document/9443178>

Arctic Sea Ice Concentration with DL: <https://www.mdpi.com/2072-4292/9/12/1305>

Melt Pond Net

Ice Deeplab

<https://www.sciencedirect.com/science/article/abs/pii/S0034425720302893>

Similarity to other task (e.g. medical images: 9,10,11,12 in River DL)

Thresholding: 8

SVM: 1-3

‘Finally, it seems that very limited existing work has been done on the application of deep learning for melt pond analysis as the only one that was found uses optical data instead of IR images.’

Model Architecture

<https://www.spiedigitallibrary.org/journals/journal-of-applied-remote-sensing/volume-16/issue-4/046514/Semantic-segmentation-of-Arctic-Sea-ice-in-summer-from-remote/10.1117/1.JRS.16.046514.short?SSO=1>

Deep Learning is widely used in Remote Sensing tasks. Most work focuses on land cover… In the area of cryospheric research,…

* The model chosen for this work is a U-Net. U-Net has been introduced in 2016 for biomedical image segmentation.
* Consists of encoder and decoder…
* Strengths: Widely used and comparison to other studies can be done. Implementations available. Relatively simple (compare to others). Won many challenges in biomedical image segmentation. Similarity to melt pond task (‘River ice images bear a significant resemblance to microscopic images of cells in the bloodstream that initially suggested the use of existing cell classification networks from medical imaging – what is different from medical images (medical imaging tasks are mainly concerned with detection and localization of specific kinds of cells rather than performing pixelwise segmentation for estimating ice concentration)
* Weaknesses: Better performing models exist. DL for Remote Sensing paper. Extensions of U-Net.
* What it means for results: Can be improved by using different model architectures, or extending layers (atrous convolutions… what for…)

“U-net is usually a good choice for small training sets in sematic segmentation. It has an encoder–decoder architecture with several shortcut connections between the two modules to incorporate multilevel features in the prediction. In this way, image details can be preserved. Its name derives from the elegant U-shaped structure displayed in Fig. 3. U-net is a small and simple network that contains nine pairs of 3×3 convolutional layers to extract features and one 1×1 convolutional layer to compress output channels. Shortcut connections are added between convolutional blocks whose indexes are summed to be 10. U-net is mostly applied in binary classification cases, while here we use a “softmax” activation in the last layer to output multiclass labels.“ (<https://ieeexplore.ieee.org/abstract/document/8601351>)

U-Net can be successfully used for RS tasks <https://scholar.google.com/scholar_lookup?title=Water+body+extraction+from+very+high-resolution+remote+sensing+imagery+using+deep+u-net+and+a+super+pixel+-based+conditional+random+field+model&author=Feng,+W.&author=Sui,+H.&author=Huang,+W.&author=Xu,+C.&author=An,+K.&publication_year=2019&journal=IEEE+Geosci.+Remote+Sens.+Lett.&volume=16&pages=618%E2%80%93622&doi=10.1109/LGRS.2018.2879492>

Limited Data

* Semantic Segmentation requires heavy data labelling, as not only classes have to be assigned to images, but in the best case pixel-wise accurate annotations need to be done.
* A major challenge for many real life and remote sensing tasks is limited availability of labelled data. As Computer Vision datasets are improved online, worldwide and… remote sensing data is often harder to access, no funding available.
* For annotating images, machine learning methods can be fused with manual.
* Research has been done to improve model performance on limited training data.

However, unlike natural RGB-images, RS-images contain several types of low-resolution objects that are irregularly shaped, which impact subsequent object classifications [[12](https://www.mdpi.com/2220-9964/9/10/601#B12-ijgi-09-00601)]. Furthermore, as RS-images are acquired from a bird’s eye perspective, the objects lie within a flat two-dimensional (2D) plane where only the top of the objects is observed [[22](https://www.mdpi.com/2220-9964/9/10/601#B22-ijgi-09-00601)]. Additionally, constructing a large-scale RS-image dataset is more difficult than using natural RGB-images, and creating data labels for RS-images obtained from various sensors is time consuming. Indeed, various errors can be introduced due to factors such as relief displacement caused by differences in elevation and shadows in the RS-images. Moreover, it can be difficult to define meaningful classes in a scene in case of numerous surface materials.

(<https://scholar.google.com/scholar_lookup?title=A+comparison+and+strategy+of+semantic+segmentation+on+remote+sensing+images&conference=Proceedings+of+the+International+Conference+on+Natural+Computation,+Fuzzy+Systems+and+Knowledge+Discovery&author=Hu,+J.&author=Li,+L.&author=Lin,+Y.&author=Wu,+F.&author=Zhao,+J.&publication_year=2019&pages=21%E2%80%9329>)

(<https://scholar.google.com/scholar_lookup?title=Semantic+Segmentation+of+Earth+Observation+Data+Using+Multimodal+and+Multi-scale+Deep+Networks&conference=Proceedings+of+the+Computer+Vision%E2%80%94ACCV&author=Audebert,+N.&author=Le+Saux,+B.&author=Lef%C3%A8vre,+S.&publication_year=2016&pages=180%E2%80%93196>)

Augmentation

* Are commonly used methods to increase the size of the dataset at hand and include rotating, scaling, shifting images. For the task at hand, … are used. Other augmentation methods, such as … do not make sense and have been disregarded because of…

Transfer Learning

Verschiedene Domänen und deren Einfluss

* To the best of the authors knowledge, there exists no labelled dataset with same task (same labels), so domain (features) are targeted to be as similar as they can 🡪 different task / same domain

<https://ieeexplore.ieee.org/abstract/document/9782149>

Strategies:

Off-the-shelf pre trained models as feature extractors (use one or more layers of a network trained on a different task as generic feature detectors, eg. Copy without final classification layer).

Additionally selectively retrain some of the previous layers by freezing certain layers while retraining

<https://arxiv.org/pdf/2204.02825.pdf>

<https://www.mdpi.com/2072-4292/14/18/4632> (for Intro about Transfer Learning / RS difficulty)

<https://arxiv.org/pdf/2205.11423.pdf> (how pretraining works in semantic segmentation)

Patch Sizes <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6033561/> (also references).

* Patch Sizes must be larger than objects to segment <https://www.youtube.com/watch?v=LM9yisNYfyw>
* Extracting smaller patches from the images can result in a larger training set whereas the receptive field gets smaller.
* Especially interesting for the task at hand, to discover which structure sizes are relevant for melt pond detection.
* Patch sizes für homogene datensatz?

**Methodology**

* Data
* Annotation
* Preprocessing
* Model Training (used implementation by…)
* Hyperparameter Tuning

Adjustments made on sm implementation: input channels. Fix input size of images (important for prediction with patches). Restructured code and training function such that hyperparameter optimization possible. Prediction function to argmax. Prediction function handle image patches (to allow for comparing results on patches of different sizes, images have been returned to their original scale (480,480).

Note on applying pretraining (imagenet) on greyscale images (<https://towardsdatascience.com/transfer-learning-on-greyscale-images-how-to-fine-tune-pretrained-models-on-black-and-white-9a5150755c7a> ):

* Copy last channel to desired channel dimension
* Modify first conv layer of model

Blogpost shows that 1) yields better results, although more resources are needed

Augmentation:

* Color transformation is not suitable (in remote sensing) because color/grayscale value is important factor in image interpretation
* Brightness and contrast transform?
* One-sample transform used only: simplicity of operation and low time cost, no geometric fusion distortions
* Geometric transformation: Rotating, scaling, flipping, shifting, cropping. Position and orientation change (makes all sense). Rotation changes the orientation of image content. Scaling enlarges or reduces image according to certain ratio (not regarded due to patches???). rotation might result in informstion loss at the image boundary 🡪 resampling method
* Sharpness transformation:
* Noise disturbance
* Random erasure
* More advanced multi-sample transforms such as mixup or cutmix are not applied because mixing multiple images carries to high risk to result in distorted shapes that are crucial for identifying the surface structures.

Patch Extraction

Patches should still be large enough to cover the characteristic shape of melt ponds. It is to determine if single ponds are sufficient or if the receptive field must be large enough to comprise interconnecting channels. In general, a trade off must be done between dataset size and spatial coverage.

Extracting smaller patches from the images can result in a larger training set whereas the receptive field gets smaller. Patch Sizes were constrained to multipliers of 32 due to the implementation used. Patches should be larger enough to cover melt ponds and their linkage and small enough to not extend a pattern. Initially, patch sizes of 32,64,256,480 have been selected. As 256 had the best initial results (see section Results), 128 has been chosen as additional patch size.

<https://www.youtube.com/watch?v=LM9yisNYfyw> !!!

Hyperparameters:

Optimizer:

Batch Size: [1, 2, 4, 8, 16]

Preprocessing: Data Normalization is a common step to preprocess data. It has been shown that steps such as scaling and normalization result in better training performance. We used Resnet34 as backbone function which includes a batch normalization layer at the beginning of the model. Resnet34 preprocessing is therefore the identity function (channels order RGB).

UNet is originally built for 3-channel data. To make UNet feasible for 1-channel data, we added a 1x1 convolution at the beginning, which increases the number of channels (WHY???)

All values are casted to float32 as Tensorflow has shown to work best with those.

480: 1, 2, 4, 8

Evaluation: IoU Score, Loss, visual comparison, comparison kmeans on visual and IR.

IS THE DATASET UNBALANCED?

**Results**

* Evaluation Metrics
* Quantitative Results
* Qualitative Results

First default training indicated that training plateau was reached at 60 epochs. 256 showed highest iou of the patch sizes. 128 was also investigated to hope for more training data while still having a receptive field high enough.

Patch size 32 was not considered for further optimization, as training results are not good enough.

Default hyperparameters were set to: 100 epochs, Adam Optimizer, 20x32, 10x64, 5x128, 3x256, 1x480 (fixed).

Hyperparameter Optimization: NAdam, RMSProp, Adam

Google: common batch sizes small dataset

The raw output is a softmax which will be probabilities. To convert this to segmentation masks corresponding to the unique classes, argmax is used.

**Discussion**

Issues in Annotating

Failure of Optical Data Accuracy

Semi-supervised approaches such as <https://ieeexplore.ieee.org/abstract/document/9460820> possible

**Conclusion**

This work tried to do this and that.

This are the results.

Note that while these improvements are promising,… limitations and future research.

Related Work:

(REVISE ALGORITHMS FOR RESULT INSPIRATION)

…, … built a random forest based approach. In IR images, boundaries are often not sharp enough to reach good results in edge-based segmentation and smaller ponds are often not detected (Figure 1).

Thielke et al implemented simple thresholding. Thresholding involves setting a threshold for the pixel brightness and classifying according to those. Thresholding requires careful manual selection of a boundary and is sensitive to environmental and seasonal changes in temperature. Even within one flight, temperatures change for different surface features and even within one image, differences can occur (Figure 2). (changing temperatures due to changing atmospheric conditions, affected by flight)

So far, no method could be developed to accurately segment melt ponds from Infrared Images. What is missing is a technique, that can take different features, characteristic temperature patterns, shapes into account. Recent advances in semantic segmentation with Neural Networks for Remote Sensing tasks 🡪 promising approach (more literature about feature extraction with Neural Networks).

Supervised methods use a set of known examples to assign a classification to unknown objects based on similarity to training data. “Can produce fixed surface type definitions, allow for more control and fine tuning of the algorithm, improve skill as more points are added to training data and allow users to choose what to classify.

Supervised methods have shown high accuracy in RS applications (Duro et al 2012).

<https://www.sciencedirect.com/science/article/abs/pii/S0165232X15001433> (Image segmentation (object grouping according to similarity of spectral and textural information; RF classifier)

<https://tc.copernicus.org/articles/6/431/2012/> (used spectral unmixing ANN for low-resolution MODIS data – maybe exclude from related work because different resolution)

<https://www.sciencedirect.com/science/article/pii/S0034425708000047?casa_token=-feoMjqZ8fsAAAAA:82EZ1MMLGxLeYd-5LuqQK8xZQk6T3tb09jjXHYLBcQqdJ6gN4taNUPSsMfk1fLD5w0I4YHhG> tschudi 2008

<https://ieeexplore.ieee.org/abstract/document/1459029> (verstehe ich net)

Rösel & Kaleschke, [2011](https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019JC015569#jgrc23831-bib-0030) <https://www.cambridge.org/core/journals/annals-of-glaciology/article/comparison-of-different-retrieval-techniques-for-melt-ponds-on-arctic-sea-ice-from-landsat-and-modis-satellite-data/3E7FBFA42BDBD8EDD84778C13FF44BE3> (PCA)