



## A Convolutional Neural Network Architecture for Sentinel-1 and AMSR2 Data Fusion

**Malmgren-Hansen, David; Pedersen, Leif Toudal; Nielsen, Allan Aasbjerg; Brandt Kreiner, Matilde; Saldo, Roberto; Skriver, Henning; Lavelle, John; Buus-Hinkler, Jørgen; Harnvig, Klaus**

*Published in:*  
IEEE Transactions on Geoscience and Remote Sensing

*Link to article, DOI:*  
[10.1109/TGRS.2020.3004539](https://doi.org/10.1109/TGRS.2020.3004539)

*Publication date:*  
2021

*Document Version*  
Peer reviewed version

[Link back to DTU Orbit](#)

*Citation (APA):*  
Malmgren-Hansen, D., Pedersen, L. T., Nielsen, A. A., Brandt Kreiner, M., Saldo, R., Skriver, H., Lavelle, J., Buus-Hinkler, J., & Harnvig, K. (2021). A Convolutional Neural Network Architecture for Sentinel-1 and AMSR2 Data Fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 59(3), 1890-1902.  
<https://doi.org/10.1109/TGRS.2020.3004539>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# A Convolutional Neural Network Architecture for Sentinel-1 and AMSR2 Data Fusion

David Malmgren-Hansen<sup>1</sup>, Leif Toudal Pedersen<sup>2</sup>, Allan Aasbjerg Nielsen<sup>1</sup>, Matilde Brandt Kreiner<sup>3</sup>, Roberto Saldo<sup>2</sup>, Henning Skriver<sup>2</sup>, John Lavelle<sup>3</sup>, Jørgen Buus-Hinkler<sup>3</sup>, and Klaus Harnvig<sup>4</sup>

<sup>1</sup>Department of Applied Mathematics and Computer Science, Technical University of Denmark

<sup>2</sup>National Space Institute, Technical University of Denmark

<sup>3</sup>Danish Meteorological Institute

<sup>4</sup>Harnvig Arctic and Maritime

**Abstract**—With a growing number of different satellite sensors, data fusion offers great potential in many applications. In this work, a Convolutional Neural Network (CNN) architecture is presented for fusing Sentinel-1 SAR imagery and the Advanced Microwave Scanning Radiometer 2 (AMSR2) data. The CNN is applied to prediction of Arctic sea ice for marine navigation and as input to sea ice forecast models. This generic model is specifically well suited for fusing data sources where the ground resolutions of the sensors differ with orders of magnitude, here 35 km x 62 km (for AMSR2, 6.9Ghz) compared to the 93 m x 87 m (for sentinel-1 IW mode). In this work, two optimization approaches are compared using the categorical cross-entropy error function in the specific application of CNN training on sea ice charts. In the first approach, concentrations are thresholded to be encoded in a standard binary fashion, and in the second, concentrations are used as the target probability directly. The second method leads to a significant improvement in  $R^2$  measured on the prediction of ice concentrations evaluated over the test set. The performance improves, both in terms of robustness to noise and alignment with mean concentrations from Ice Analysts in the validation data, and an  $R^2$  value of 0.89 is achieved over the independent test set. It can be concluded that CNNs are suitable for multi-sensor fusion even with sensors that differ in resolutions by large factors, such as in the case of Sentinel-1 SAR and AMSR2.

**Index Terms**—Synthetic Aperture Radar data, Microwave Radiometry, Cryosphere

## I. INTRODUCTION

**D**ATA Fusion of satellite sensors has a growing potential to improve applications in Earth Observation (EO) with the growing number of EO satellites being launched. Observation of the Earth with a satellite sensor is often an indirect measurement of the underlying physical/biological process one wishes to estimate, e.g. the satellite sensor measures microwave radiation while we want to estimate ice concentration. The idea of combining measurements from multiple sensors can offer new relationships to the physical process and strengthen the information retrieval.

In this work, a new architecture of a Convolutional Neural Network (CNN) is presented to facilitate the fusion of sensors with very large ground resolution difference. The architecture

differs from other fusion networks in that all sources are not simply re-sampled and stacked into an extension of image bands before provided as input. Instead, the network saves a large amount of computations by concatenating the low resolution data with the CNN extracted image features at a deeper layer in the network.

While the proposed network architecture is generally useful when sensors have large resolution differences, it is applied here to the problem of generating Arctic sea ice maps from Synthetic Aperture Radar (SAR) images combined with microwave radiometer data. In the context of sea ice mapping, SAR images show the spatial patterns formed by sea ice on the ocean in high resolution (<100 m) well, but its backscatter values do not always distinguish between open sea in windy conditions and various ice surfaces. On the other hand, radiometers such as the Advanced Microwave Scanning Radiometer 2 (AMSR2) that measures brightness temperatures show a very distinct difference between ocean and ice relatively independent of the wind conditions, but at a rather coarse resolution. The potential in combining the information from the two sensors is therefore large and will be explored in this work.

SAR satellite images have been used extensively for producing sea ice charts in support for Arctic navigation. However, due to ambiguities in the relationship between SAR backscatter and ice conditions (different ice types and concentrations as well as different wind conditions can have the same backscatter signature) the process of producing ice charts consists of manual interpretation of the satellite data taking into account also the texture patterns of the ice and ocean in the SAR images. The process is labor-intensive and time-consuming, and thus, the amount of ice charts that can be produced from SAR images on a given day is limited. The long time (often 6-12 hours or more) from data acquisition by the satellite to delivery of the ice chart to users also reduces the value of the chart information, especially in dynamic sea ice regions. Automatically generated high resolution sea ice maps can potentially lead to increased use of satellite imagery in ice charting and navigation by producing ice products faster and more frequently, covering larger geographical areas.

The design of an automatic and robust sea ice classification scheme has been studied for many years, [1]. Recent

approaches to this issue using CNN (Convolutional Neural Networks, as an image segmentation technique) show promising results, [2]. In this paper we explore the CNN framework for generating ice products and expand it by introducing a new satellite sensor fusion technique that allows the CNN to exploit Microwave Radiometer (MWR) data together with SAR images for improved end results.

Contributions of this paper are, 1) a general CNN architecture for combining advantages of data acquired using two or more sensors of very different ground resolution, 2) the application of the methodology to automatic generation of sea ice maps and 3) improvements in CNN training by using “soft” probabilities in applications where these are available, such as the Sea Ice Concentrations (SICs) used here. In terms of the CNN’s architecture, the proposed method provides several new concepts. It concatenates very low resolution sensor data at a deeper layer in the CNN model in order to save convolutions, i.e. computations, rather than stacking all data as an extended band image. With this concept, image features are extracted from the high resolution sensor and combined with information at a lower scale to improve the pixel-wise classification, i.e. semantic segmentation. The concatenation method also enables inclusion of additional input data that could provide information of the problem at hand, such as wind data in ice charting.

To show the value of the method it is applied to the case of segmenting sea ice from SAR images. To the authors’ knowledge it is the largest scale experiment in sea ice mapping to date with data from 912 Sentinel-1 scenes distributed over the largest geographical area, covering sea ice conditions in all seasonal variability. It is further suggested to improve the segmentation performance by training the CNN with ice chart concentration as the categorical probability, referred to as “soft” probability, in the loss function rather than binary zero/one encoded target values. The most significant contribution of the study on sea ice mapping is the combination of a SAR system on image level, with radiometer data at meta level in order to solve the ambiguities in SAR data when distinguishing open sea from ice in certain ice and weather conditions.

## II. DATA

A dataset was prepared by selecting a large set of regional ice charts (also commonly referred to as Image Analysis, e.g. in [2]) that were manually produced by the ice analysts of the Greenland Ice Service at the Danish Meteorological Institute (DMI) and based primarily on Sentinel-1 SAR imagery. SAR imagery has proven very suitable for Arctic sea ice monitoring due to the radar sensor’s capability to see through clouds and in polar darkness. The 912 ice charts and corresponding SAR imagery cover the period from the availability of the Sentinel-1A sensor data in November 2014 and Sentinel-1B data from September 2016, until December 2017, with samples over all months, see Table I. It should be noted that due to the way ice charts are drawn they do not necessarily cover a full Sentinel-1 scene. The amount of Sentinel-1 pixels that are covered by an ice chart polygon corresponds to approximately 300

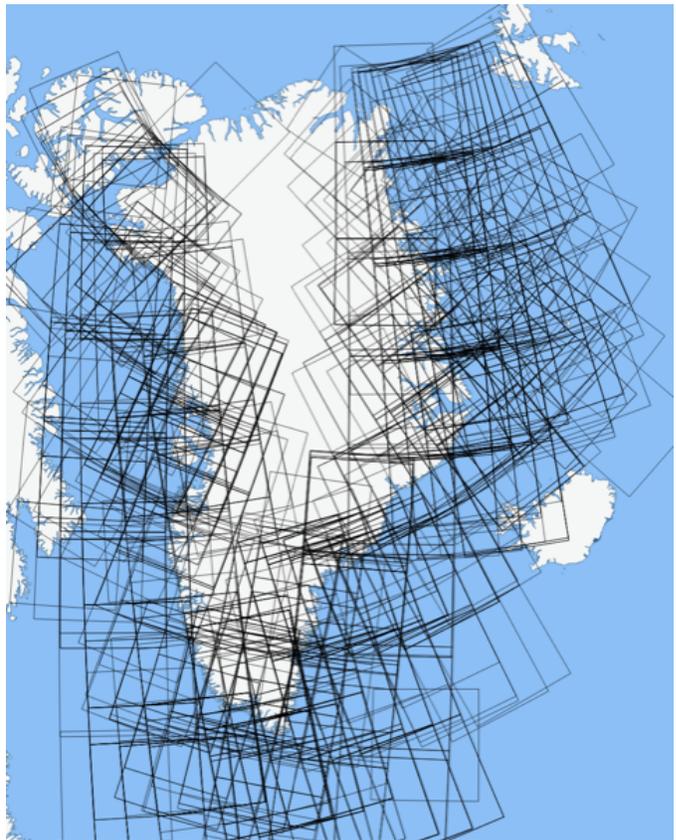


Fig. 1: Outlines of the 912 Sentinel-1 scenes in the ASIP dataset covering most of Greenland’s coast.

TABLE I: Number of scenes per month in the dataset.

Month	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
# of scenes	54	38	67	62	74	93	93	115	95	78	70	73

full scenes. Outlines of scenes can be seen plotted on a map in Figure 1. The individual ice charts cover different areas (depending on users and season) but the dataset spans all of Greenland’s coast, apart from the northernmost regions, and with denser coverage in Southern Greenland. The combined size of the dataset is 156GB and it is available for download at [3]. For the studies in this paper data from 88 scenes, approximately 10%, were kept aside from the training process as a testset. It was deliberately chosen to split the data on scene-level rather than smaller samples within-scene in order to separate training- and test-set by either time or geographical region while still making sure that both datasets include all seasons and regions.

### A. Sentinel-1 SAR

We use C-band SAR data from the Copernicus Sentinel-1 A and B satellites. The data utilized are medium resolution level 1 ground range detected scenes (GRDM) in Extra Wide swath mode (EW) with a resolution of 93 m x 87 m (range by azimuth) and a pixel spacing of 40 m x 40 m. EW GRDM has been chosen because it is the most frequently used mode in marine polar regions with the most favorable combination of coverage and resolution. The Sentinel-1 SAR

TABLE II: Resolution of AMSR2 at different frequencies, [4]. [IFOV] Instantaneous Field Of View. P. dist. is the distance between AMSR2 measurement points.

Freq. [GHz]	6.925	7.3	10.65	18.7	23.8	36.5	89.0
IFOV [km]	35x62	35x62	24x42	14x22	11x19	7x12	3x5
P. dist. [km]	10x10	10x10	10x10	10x10	10x10	10x10	5x5

data are used in their original GRDM geometry in order to avoid resampling the data with potential effects on the image statistics and texture. The data are calibrated to backscatter and converted to decibel (dB) before use. To have a sufficient number of co-located ice charts we include data from the early period of Sentinel-1 so the HV channel data suffer from quite severe subswath striping and scalloping effects. More recent data (especially after June 2019) include corrections for these effects and will be used in the future when sufficient ice charts become available. The Sentinel-1 SAR data are received in near real time through the Copernicus Marine Environment Monitoring Service (CMEMS) of the European Union. The data are also available from the Open Access Data Hub at: <https://scihub.copernicus.eu/dhus/#/home>.

### B. AMSR2

The AMSR2 data is processed by resampling swaths to the coordinates of every 50<sup>th</sup> pixel in the Sentinel-1 image (corresponding to every 2 km). This results in different oversampling of the AMSR2 frequency bands since each band has a different size footprint, see Table II. In an operational scenario the Sentinel-1 scene is typically available for the DMI Ice Service 1-2 hours after its acquisition as opposed to AMSR2 which is available in near real time. This has been taken into account when searching for AMSR2 swaths to match with the Sentinel-1 scenes for the dataset by searching for the nearest AMSR2 match in time. The AMSR2 data are provided by The Japan Aerospace Exploration Agency (JAXA). We use L1B data downloaded from <ftp://ftp.gportal.jaxa.jp>. AMSR2 is an instrument on board JAXA’s GCOM-W satellite (Global Change Observation Mission 1st-Water).

For each Sentinel-1 SAR image in the dataset, a matching AMSR2 source is produced by finding the AMSR2 acquisitions which, in the seven-hour window preceding the timestamp of the Sentinel-1 pass and 2 hours following, intersects with the Sentinel-1 image. Each of the AMSR2 passes in this temporal window, is then resampled onto the subsampled Sentinel-1 pixel coordinates using Gaussian weighted interpolation with the Python *pyresample* library. This results in a time series of resampled AMSR2 data layers, some of which have missing pixel values, due to the Sentinel-1 and AMSR2 swaths not overlapping completely. From this co-located time series, the pixels without missing values and with the smallest time difference to the Sentinel-1 image are selected across the time dimension. In some cases, this results in a mosaic AMSR2 file, where pixels are obtained from two or more swaths. The chosen time window is sufficient to ensure that there are no missing values in this resampled AMSR2 swath, provided that there were no AMSR2 data outages. The median time delay between the nearest AMSR2 and the 900+ Sentinel-

1 scenes in the dataset is 4h45m and all scenes are matched within 6h40m time difference.

### C. Ice Charts

Manual ice charting from multi-sensor satellite data analysis has for many years been the common method at the national ice centers around the world for producing sea ice information for marine safety. It should be noted that different ice services have slightly different definitions of the ice chart products and may also refer to them as Ice Analysis Chart or Image Analysis Chart. The ice charts in the dataset used here are from the archive of the operational sea ice service at the DMI. DMI has produced sea ice concentration charts for the waters around Greenland for more than 50 years. Today, DMI Ice Analysts primarily use Copernicus Sentinel-1 SAR imagery. The Sentinel-1 imagery used for ice charting are primarily in EW-mode and dual polarization (HH+HV). Auxiliary satellite data such as optical imagery, thermal-infrared and passive microwave radiometer data, as well as numerical weather prediction data are used when available to support the analysis. The ice chart coverage and frequency depends on season, ship routes and data availability. Normally, one ice chart is produced on the basis of one Sentinel-1 image and is thus a snapshot of the ice conditions at the time of the SAR image acquisition. The ice analysts use the latest available Sentinel-1 SAR image and carry out a detailed manual interpretation by drawing the ice chart in an ArcGIS production system. An ice chart consists of manually drawn polygons of fairly homogeneous ice conditions given as a concentration of ice inside the polygon from 0-100%, where 100% ice is completely covered ocean and 0% ice means open water. The estimates of ice concentration in the charts are based on the subjective evaluation by the analyst and individual ice charts or polygons have no associated uncertainty. Analysts pay particular attention to regions near the ice edge because the characteristics and extent of ice in the marginal ice zone are important for operations taking place within or near that region. Conversely, the DMI analysts generally do not characterize the inner ice zone with as much attention to detail, because most of the time there are no supported navigation there. It is important to notice that the relative accuracy and level of analysis detail varies. Studies of ice charts from different ice centers covering the same geographical region show relatively large (up to 20%) variability in ice concentrations, [5]. The original ice charts can be found in the ice chart archive at the DMI homepage. The ice charts follow the World Meteorological Organization’s (WMO) code for sea ice, where the ice concentration classes are defined by concentration intervals, spanning over 10 to 30%. When converting the original ice chart format into training data for this study, these concentration intervals are converted to their mean ice concentration. The range between the concentration values are therefore 5% (eg. 0%, 5%, 10%, 15%,..., up to 100%). The ice concentrations are seldom homogeneously distributed inside the drawn polygons, meaning that the concentration of a smaller subset of the polygon, e.g. a 40 m x 40 m pixel does not have to have the same concentration as the polygon it lies within. The

average concentrations of subsets are considered correct for the polygon but the distribution of ice that averages to the given percentage can vary a lot. Ice charts are produced for areas where there is maritime traffic and since the majority of the ships in Greenland waters try to avoid sailing in sea ice, the ice maps are focused on showing areas of open water in between sea ice. In eastern and southern Greenland sea ice drift along the coast and the ice charts covering these regions often have large open water polygons and a narrower area of sea ice. These ice charting practices lead to the dataset having an over-representation of open water samples of  $\sim 70\%$  and the remaining  $\sim 30\%$  being ice samples distributed over other concentration values,  $> 0\%$ .

Due to the relatively poor quality of the Greenland coastline data used in the ice charting software (with offset and missing islands) a land-mask buffer-zone of 10 km is applied to the ice chart 2D variable in the training and test data. However, the final CNN model is capable of producing results much closer to the coast.

### III. RELATED WORK

#### A. Satellite Sensor Fusion Techniques

The overall objective of satellite sensor data fusion is combining advantages from two or more sensors for improving derived data-products, compared to using a single sensor only. With a vast amount of EO sensors operating today, the potential for development in sensor fusion techniques is larger than ever. The fusion of data from satellite sensors has been used to address a wide range of problems through time, and numerous techniques exist to accomplish this. For an overview of the field, we refer to the following review papers, [6]–[8]. In satellite sensor fusion, one distinguishes between three overall types, *Pixel Level*, *Feature Level* and *Decision Level*, [6].

In *Pixel Level* fusion, two or more co-registered sources, typically images, are combined to reveal more details than what was available from a single source. An example is pan-sharpening of multi-spectral images where the goal is to achieve as high spatial resolution as the pan-chromatic band for all multi-spectral bands in the sensor, e.g. as in [9].

*Feature Level* fusion is the concept of first extracting features from different sources and then using them as inputs to Statistical Learning models. One example is land cover classification where fusion of SAR and optical images can lead to improved results, [10].

In *Decision Level* fusion, each source of data has undergone processing and preliminary analysis, such as classification. An example of multi-sensor fusion at Decision Level is shown in [11] where several pre-classified data sources are combined and compared with different weighting schemes.

Some of the challenges in multi-sensor fusion is the spatial co-alignment of the data sources as well as the potential temporal gap between the recordings. Another challenge is how to fuse the data from sensors with very large difference in spatial resolution. We address this last challenge by utilizing the CNN model’s architecture which handles this in a computationally elegant way. We propose a Convolutional Neural

Network (CNN) architecture capable of handling data from different sources where the spatial sampling resolution is many factors higher in one sensor. For the case of the Sentinel-1 and AMSR2 fusion, the resolutions differ with magnitudes.

The proposed model in this paper takes Sentinel-1 and AMSR2 as input. Our model does a number of filter/convolution operations on the SAR image before injecting the MWR data and the approach can therefore be seen as a *Feature Level* fusion strategy. The proposed model is described further in Section IV and illustrated in Figure 2.

#### B. CNN Segmentation

With the increased use of deep learning in a wide range of areas in recent years one of the promising applications of CNNs have been obtaining improved for image segmentation (here defined as splitting an image into “segment” by pixel-wise classification), [12]. The convolutional layers of a CNN that applies a number of filters to an input image (or other data types) are well suited to learn how to extract useful features for classifying information, either at image-level or pixel-level. The filters in a network is ordered in a hierarchical fashion by stacked layers each consisting of a filter bank. For classification, information from all pixels is reduced to a single label, and layers are therefore interleaved with subsampling schemes that reduces the spatial extent and makes the model invariant to translations of the objects in the image. In the context of CNN image segmentation, the use of subsampling comes with a trade-off on the models ability to create sharp edges around objects in the output segmentation map. Currently, two main concepts of CNN architecture is used for image segmentation across many fields of use. The first, named *U-Net*, is based on the work in [13] which was further refined with “skip-connections” in [14]. In the *U-Net* architecture convolutional layers and subsampling schemes are applied to an input image similar to traditional CNN architectures for classification tasks. The subsampling reduces the spatial dimension and allows for more convolutional filters to be trained using the same or fewer computations. Another advantage of subsampling in terms of segmentation is the relative increase of the “receptive field”, i.e. is the amount of neighborhood pixels in the input image considered in the classification of each pixel in the output. When applying subsampling, such as a 2x2 Max-pooling function, the following convolutional filters consider a factor of 4 times larger area from the input image, i.e. the receptive field grows by a factor of 4. The size of the “receptive field” is considered important to capture the spatial semantics in the segmentation tasks. In order to go from a low spatial dimension back to a segmentation map with the same spatial dimension as the input image, *U-Net* applies upsampling operations interleaved with convolutional layers. However, subsampling corresponds to low pass filtering, thus high-frequency image details are lost in this process. For retaining sharper edges, i.e. high frequency information in the final output segmentation map, “skip connections” before each subsampling step is concatenated to the corresponding upsampling step that has the same spatial dimension. *U-Net* was originally developed for a medical image application, but has now proved useful in a far range of fields and applications.

As opposed to the U-Net, the second concept of CNN segmentation architectures, introduced in, [15], [16] does not apply any subsampling operations. This is to avoid the low-pass filtering property, but the high spatial dimension throughout the layers leads to a relatively higher computational cost on average per layer. To retain a "receptive field" similar to that of a CNN that includes subsampling, *Atrous* convolutions are applied. In an *Atrous* convolution, the kernel coefficients are separated with a certain number of zeros or holes (from French *à trous*: with holes) in order to have a computational low cost kernel that extends over a larger area in the input. The *Atrous* convolution thereby increases the receptive field and eliminates the need for spatial subsampling for incorporating image features at larger scales. *Atrous* filters were originally introduced by Holschneider et al. for scale dependent wavelet transforms in [17].

In [15] the authors evaluated four different strategies for incorporating scale space in a CNN model and concluded that a strategy they named Atrous Spatial Pyramid Pooling (ASPP) performs best. An ASPP CNN model applies first regular convolutional layers and then four parallel *Atrous* convolutions with different dilation rates, i.e. different spacing between coefficients. The output of the four *Atrous* layers are concatenated and a 1x1 convolution is applied to produce the output segmentation map. The model achieves an image feature extraction at different scales, and is known to be improving context modelling in image analysis, [18].

Except for obvious numerical and computational differences regarding the use of subsampling in a CNN or not, the largest conceptual difference between the two concepts lies in how they handle scale dependent information. In a U-Net style network, features from different scales are added from coarse scales to finer grained information in the stepwise upsampling process of the so called decoder part of the network. In the *Atrous* style network, scale information can be handled either by applying the *Atrous* operator throughout a network on sequentially stacked convolutional layers or alternatively by an ASPP concatenation module. The latter is chosen in this work due to the concatenation layer allowing for a direct extension beneficial for our fusion methodology, which is further explained in Section IV.

### C. Automating Sea Ice Products

Previous literature on automatic extraction of sea ice information from satellite data takes different approaches to obtaining training and validation data. The most common automatic product is derived from microwave radiometer data alone, and very high accuracy can be achieved [19]. However, the coarse resolution of the microwave radiometer SICs is a major drawback when it comes to usage for navigation where more details are desired. Until recently, attempts to automatically derive sea ice concentration from SAR data have been using pre-compiled texture measures such as based on gray-level co-occurrence matrix methods [2]. Often, ice charts produced by expert Ice Analysts are the only source of the target information, and although as described in Section II-C these products contain uncertainties in terms of

inter/intra-operator variability they serve as the only way of data annotation. Further, in [20] it is described how some ice services deviate from the World Meteorological Organization's (WMO) guidelines. As opposed to the WMO guidelines, the approach of the Finnish Meteorological Institute described in [20] attempts to increase the consistency of the labelling process of polygons in ice charts by always assigning the same information on ice. Wang et al. [2], turns the problem into a CNN least-squares regression task over SICs with RADARSAT-2 images as input. The authors work on patches of 45x45 pixels that have a pixel spacing of 400 m giving a patch size of 18 km. When evaluating new scenes, the network is applied on a sliding window manner over all pixel positions to yield a higher output resolution and a smooth looking prediction. When doing regression over SICs, the final resolution cannot be too fine as it will introduce a large amount of representation errors. This is due to the fact that ice charts consist of large polygons, each labelled with its average ice concentration only. If one looks at a small sample from a polygon, the true ice concentration can differ quite a lot from the polygon average SIC value compared to the ice chart.

Wang et al. and Karvonen, [2], [20], use two different modalities of Neural Networks, CNN and Multi Layer Perceptron respectively. They both use the least square error loss function to optimize network parameters, which corresponds to a maximum likelihood optimization of NNs given a normal distributed prediction error [21]. Whether the Gaussian assumption is valid for SICs is not reported but both papers reach convincing results in terms of overall SIC maps. [20] states that SIC levels below 25% have high uncertainty and finds that SICs below 25% often are open water. Therefore, [20] suggests thresholding of SICs < 25% and set everything below to 0.

As an alternative to a regression task, one could consider predicting an Ice Mask with a classification loss function and maximize categorical probabilities of ice vs. water. This approach is adopted by Leigh et al. and Zakhvatkina et al., [22]–[24]. In [22]–[24], the authors developed algorithms that classify Ice/Water based on Support Vector Machines. The final product resolution in [22]–[24] range between 200 m and 16 km.

## IV. CNN ARCHITECTURE

The design-goal of the proposed CNN architecture is to incorporate certain capabilities. The first, is the fusion of two satellite data sources, the Sentinel-1 SAR data and AMSR2 MWR measurements, both measuring distinctively different electromagnetic properties on Earth at very different spatial resolutions. The second, is to capture spatial patterns, i.e. image features, in the SAR image at different scales in order to retain very high resolution, i.e. sharp edges, as well as mapping the large scale features that distinguish ice from open water in windy conditions. The third, is to retain the absolute backscatter value in the SAR images while letting the CNN extract optimal feature descriptors. When using Batch Normalization layers in the CNN, where a standard normalization is applied between layers, the scaling is performed on a small batch of

data and the relative backscatter values between batches might be shifted. Batch Normalization is a regularization technique that is known to accelerate the training and allow for deeper networks since gradients are controlled by scaling, [25]. By letting the output of the second convolutional layer, which has not undergone batch normalization yet, circumvent the rest of the layers and adding it to the concatenation just before the output layer, we ensure these backscatter values has not been scaled in a batch-wise manner. If any information on the sea ice problem lies in the absolute value of the SAR signal, we ensure this is kept in the classifying output layer. According to [26], the simplest way of performing data fusion with CNNs is by stacking the different sources and thereby extending e.g. the RGB image to more channels, as done in [27] with RGB data and a Digital Elevation Model. In the case of very different spatial resolution as with Sentinel-1 SAR and AMSR2 microwave radiometry, stacking the sources leads to an enormous amount of unnecessary redundant convolutions performed on many channels, as data must be re-sampled to the highest resolution before they are fed to the network. The architecture used in this work therefore injects the low resolution AMSR2 data at a later stage in the CNN, as seen in Figure 2, and in this way it fuses the sources intelligently. Sentinel-1 SAR input data are therefore handled separately with 6 pairwise 3x3 kernel convolutional layers with Batch Normalization and Dropout in between. To model the image data at different scales the ASPP module is adopted in the second to last layer of the model. The ASPP module extracts features at four different scales by performing *Atrous* convolutions with four different dilation rates and concatenating the outputs. Thereby features are extracted at both coarse and finer levels and it is up to the last layer to mix these in a way that is optimal for the optimization problem. Average pooling layers, i.e. 2D moving average filters, were applied with window size equal to each of the four dilation rates. This was necessary to avoid artifacts in the predictions due to the “hole” between filter coefficients in an *Atrous* filter.

The input size of the Sentinel-1 image patch, as well as the output prediction mask, is 300x300 pixels, i.e. 12 km x 12 km, of each HH and HV channel. The corresponding AMSR2 resampled values is 6x6 pixels of 14 channels consisting of seven frequencies in two polarizations brightness temperatures. After training, the CNN on 282 948 training samples cut out without overlap from 824 SAR scenes in the training set, the model can be applied on a new unseen SAR/AMSR2 scene in sliding window manner. Each window is overlapped with 50 pixels and the 50 pixel edge around each prediction output is discarded to avoid predictions with padded values. We trained all models with the Adam optimizer and left the values of its parameters as advised in [28]. Training was performed for a fixed 80 epochs for all models as this generally ensured a good convergence was reached. Code for the model definition can be found at [29].

#### A. Optimization Strategies

The choice of final product resolution and whether to model the problem as regression over SIC or classification

of Ice/Water is connected to the end-user need. Typically, SICs are provided over large polygons that will indicate zones where ships of certain ice-grade can or cannot navigate, but these products are not always sufficient. In terms of Ice/Water products these would often be needed in much higher resolution as it might be used e.g. by fishing vessels that want to navigate as close to the ice-edge as possible without encountering dense ice. As our user-survey (report available in Danish at [30]) showed a demand for very high resolution products our method is based on a CNN model where pre-subsampling is not necessary. This allows for experiments with final product resolution as high as possible and for the model to be transferred to even higher resolution sensor data if this would become available in the future. In terms of training a CNN to classify pixels, the categorical cross-entropy loss function for maximum likelihood classification is used as written below,

$$E(\mathbf{w}) = \sum_i t_i \log(y(\mathbf{w}, x_i)) \quad (1)$$

Here,  $y(\mathbf{w}, x_i)$  is the output of the CNN for the  $i$ 'th sample given the weights  $\mathbf{w}$  and  $t_i$  is the target for  $i$ 'th sample. The most common way to encode  $t_i$  is  $[0; 1]$  where 1 indicates ice and 0 open water. For the ice charts to provide this encoding they must first be transformed to zeros and ones by applying a threshold at a certain SIC value, as in [22], [23]. This will introduce some amount of label errors as a pixel within a polygon in the chart might be open-water even though the mean ice-concentration is high or vice versa. Neural Networks are known to cope well with some amount of label errors due to the stochastic nature of the training process, [31]. Alternatively, we propose to use the SICs as the “true” probabilities of finding an ice-pixel within an ice chart polygon. Under the assumption that most pixels at a scale of the sensor resolution will be either completely ice or open water, the concentration of a polygon corresponds to the probability of a given pixel within that polygon being ice. Additionally, we also ensure to reduce the effect of label-errors as we “punish” the network less for predicting water-pixels in a e.g. 80% ice zone. The concept of using probabilities as “soft targets” was also used in [32] to *distilling* the knowledge learned by a large model on a training set into a smaller model, by using the larger models output probabilities. According to [32], the ratio between different class probabilities are said to contain richer information about the problem at hand than binary labels. In Section V we shall investigate the model’s ability to align its predictions with the Analysts ice charts using either the binary encoding of ice/water or the SIC target probabilities.

## V. RESULTS

One might consider different metrics for evaluating models depending of its use for properly addressing the variations in the use of sea ice products. For the majority of users, [30], in arctic marine navigation the main concern is to avoid areas with ice. Another application could be to monitor ice quantities in climate studies or use them as input to forecast models where the main focus would be accurate concentration values.

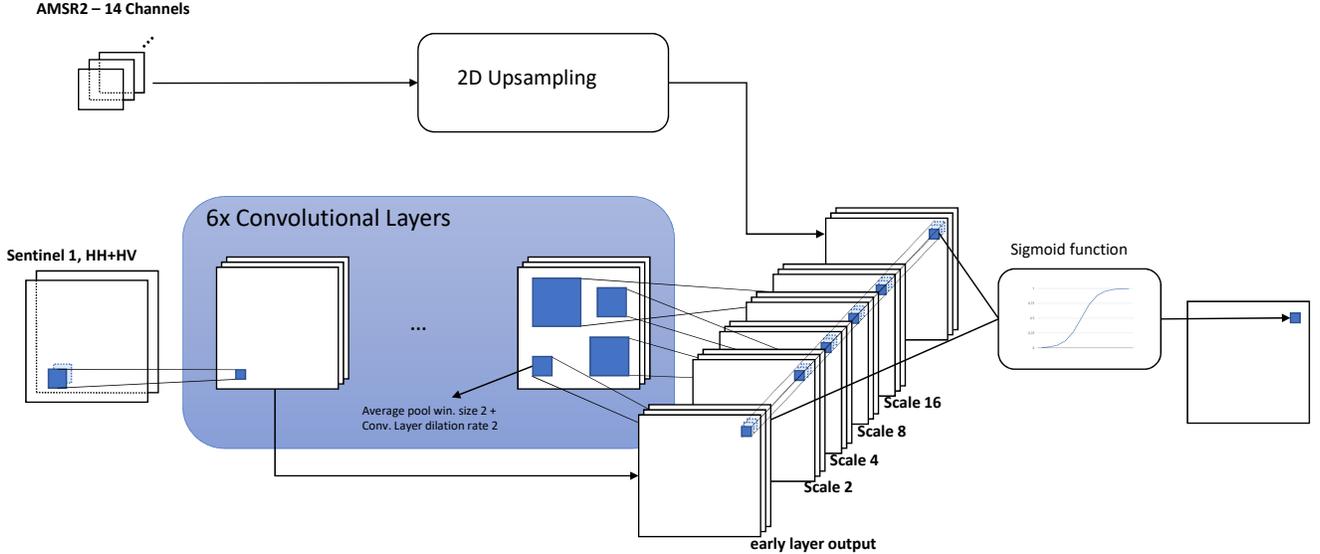


Fig. 2: CNN architecture. The 2D upsampling performed on the AMSR2 data inside the CNN is done in order to implement the last layer as 1x1 convolutional filter kernels.

Predicting the ice-zone, defined as the boundary between open water and an area with any concentration of ice, is appropriate for the users that wish to avoid sailing in any ice. We therefore evaluate the models with, 1) a pixel-wise accuracy measure in terms of ice/no-ice zones with ice charts thresholded at 10% alongside, 2) a mean predicted SIC vs. analyst estimated SIC as a mean over all polygons within a certain concentration value. Evaluation 1) is shown in Table III as the balanced accuracy, given by the mean of the sensitivity,  $SE$ , and specificity,  $SP$ , with  $SE = TP/(TP + FN)$ ,  $SP = TN/(TN + FP)$  where  $TP$  is number of True-Positives,  $TN$  True-Negatives,  $FP$  False-Positives and  $FN$  False-Negatives. The first sub-table provides the performance of two models trained on binary ice masks extracted from thresholding the DMI ice charts at a 10% SIC level. The pixel-wise probabilities from the two models in the first sub-table have been thresholded at 50% to achieve a binary mask to compare with, since 50% is the middle of the 0/1 labels used for their training. In the second sub-table, the two models have been trained to reflect SIC values from the ice charts as the probability ( $t_i$ ) in Equation 1. These models are trained to provide probabilities that reflect the SIC values. As opposed to a model trained on binary ice labels from thresholding charts

at 10% which should have a 50% probability threshold, they have to be thresholded at 10% to give a suitable comparison. By testing at different output probability thresholds for both modelling approaches we can confirm that the thresholds are optimal as seen in Table III. The results in Table III show that we add substantial information to the model when training with SIC values as probabilities rather than binarizing the problem and training with 0/1 labels. Soft targets such as the SIC values here are improving the results of CNN training in line with findings in [32].

The pixel-wise accuracy in Table III is showing the performance on predicting whether a pixel belongs to an ice zone or open-water. Alternatively, one might be interested in testing the models' overall alignment with analyst assessed SIC values as a performance measure for replicating the analyst's task. This can be done by comparing the mean predictions within each of the test set polygons, as the analyst's estimate also is an average estimation. The polygon average of ice according to the CNN output can be calculated in two ways. Either, by averaging the probabilities from the CNN or, by counting the fraction of pixels assigned a probability higher than 50%. The latter approach has been used for a comparison of the models' alignment with ice chart from analysts in Figures 3 and 4.

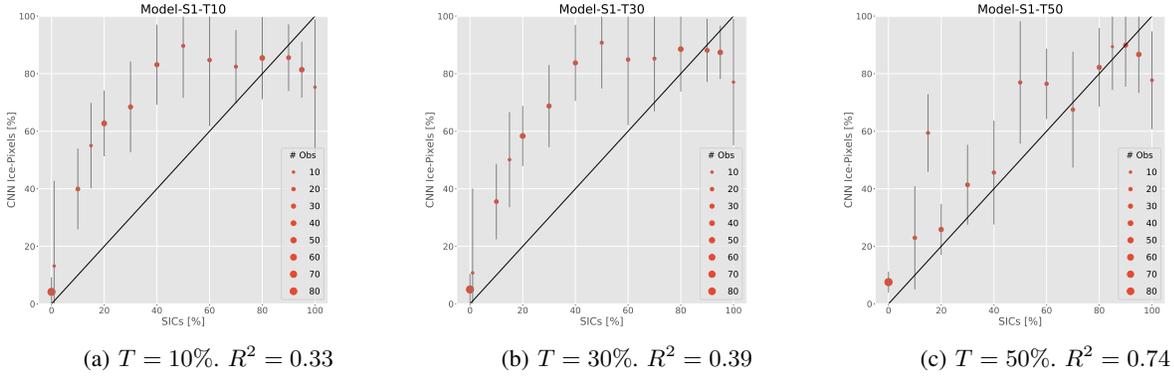


Fig. 3: Summary of testing against expert labelled ice concentrations when the CNN model is trained on i.e. labels from the ice charts being thresholded at a given concentration level. Y-axis is the CNN predictions, fraction of pixels classified as ice within each polygon, X-axis is the SIC value from the ice chart. Ice concentration levels with less than five examples in the test set have been excluded. The red dots is mean over all polygons in the testset and the grey line is the standard error.

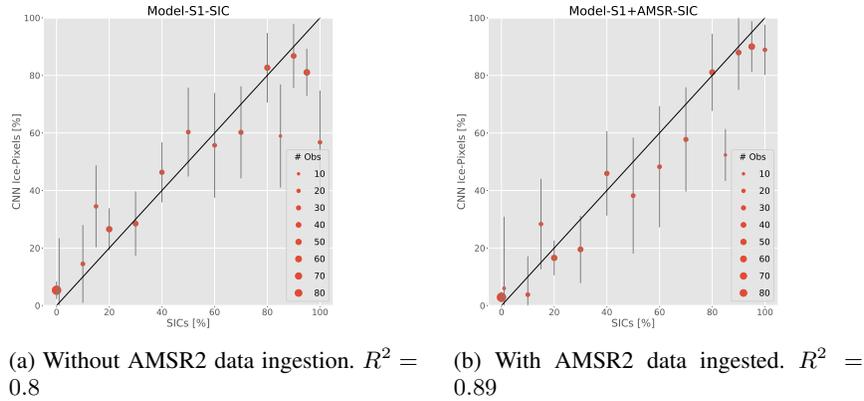


Fig. 4: Summary of testing against expert labelled ice concentrations when the CNN model is trained on categorical probabilities being set to the ice concentration level from the ice charts. Y-axis is the CNN predictions, fraction of pixels classified as ice within each polygon, X-axis is the SIC value from the ice chart. The red dots is mean over all polygons in the testset and the grey line is the standard error.

TABLE III: Pixel wise accuracy of the four models with the ice chart being thresholded at 10% concentrations to convert them to binary labels for ice/water. Evaluated with the test set of 28'651 image-windows of 300x300 pixels.  $T_p$  is the threshold on the CNN output probability at which we consider a prediction ice.

Name	Model-S1-T10	Model-S1+AMSR-T10
Input	S1	S1 + AMSR2
Opt. Method	ice charts thresholded at 10%	
Acc., $T_p = 10\%$	86.82%	89.85%
Acc., $T_p = 30\%$	91.24%	94.76%
Acc., $T_p = 50\%$	<b>91.81%</b>	<b>95.29%</b>
Acc., $T_p = 70\%$	91.47%	94.91%
Acc., $T_p = 90\%$	89.07%	92.93%

Name	Model-S1-SIC	Model-S1+AMSR-SIC
Input	S1	S1 + AMSR2
Opt. Method	SIC used as categorical probability	
Acc., $T_p = 10\%$	<b>90.64%</b>	<b>94.26%</b>
Acc., $T_p = 30\%$	88.5%	89.46%
Acc., $T_p = 50\%$	81.39%	82.40%
Acc., $T_p = 70\%$	73.51%	75.85%
Acc., $T_p = 90\%$	55.89%	64.27%

Figure 3 shows the difference between model predictions when we change the target ice mask by thresholding at different SIC values. It is not surprising that thresholding at 50% makes the models prediction align better with the ice charts than 10% but in Figure 4 it is seen that far better results are achieved when training with SIC values as targets in terms of the  $R^2$  measure. Further, it is also advantageous to add AMSR2 data as seen on Figure 4b, where higher ice concentrations (>80%) are more accurate. The red points in Figure 3 and Figure 4 are averages of all polygons in the test set with the size of the dot corresponding to the number of samples,  $N$ , used to calculate the mean. The black line through each point is the standard error calculated as  $1.96 * \sigma / np.sqrt(N)$  where  $\sigma$  is the standard deviation and  $N$  is the number of unique ice concentrations in each scene.

Visual examples of the four CNN model variants can be seen in Figure 5 and 6. It is seen that both inclusion of AMSR2 data and optimizing the CNN with SIC concentrations as cross-entropy probabilities lead to smoother predictions. The noise seen in the predictions in Figure 5 comes from noise in the first subswath of Sentinel-1's TOPSAR mode. This problem

has been reported before and has recently been addressed in [33]. The noise is primarily seen inside the first subswath, near the boundary between first and second subswath and is occasionally also visible between the other subswaths. In the most recent version of the Sentinel-1 processor (IPF 3.1) the problem has been reduced significantly, but this is only available for data issued after June 6, 2019. The output probabilities on Figure 5 (d) where the model is trained with SIC as a soft probability exhibit visual smoother ice maps that reflects the true ice concentration. Comparing 5 (d) with (b) where binary labels was used clearly show the differences between the two training methods that also are reflected on the plot in Figure 3 and 4. By measuring accuracy of polygons individually, scenes with poorer performance can be found such as the one shown in Figure 7. Other than a similar noise problem as Figure 6, Figure 7 also shows a problem (marked by a red circle in the predictions) on a homogeneous ice surface where the center is classified as water. This is likely caused by the ambiguity of the SAR backscatter values on homogeneous ice and open water in windy conditions combined with the CNN not having a sufficiently large Field-of-View to capture the ice edge when classifying pixels in the middle of the ice-patch. The CNN is not able to completely suppress this noise in predictions but generally copes better with it when AMSR2 is included and SICs are used as target probabilities, see Figure 5d. Final operational products can be designed in several ways on the basis of the CNN outputs. One such example is shown on Figure 8 where the outline of the ice edge (edge between ice in zones and open water). This has been done by applying a 3x3 average filter on thresholded output of the CNN and extracting the vector outline geometry in the QGIS software package.

## VI. DISCUSSION AND CONCLUSION

This study presents the first CNN model for combining low resolution EO data, here AMSR2, with data that is significantly higher in resolution, here Sentinel-1 SAR. The model learns to extract image features from Sentinel-1 SAR imagery and combines it with the low resolution data from AMSR2 by linear projection similar to a logistic regression. Two modelling schemes are proposed, one where the ice charts are thresholded and used as 0/1 target variables similar to how most segmentation and classification tasks are solved. This showed to be inferior in modelling the ice charts presented here as opposed to the second method where we used the ice concentration values directly in the cross-entropy error function we minimize during training. The binary approach of optimization might still have value in applications where a conservative binary ice mask is sufficient for satisfying user needs though. It can be concluded from the experiments shown here that both the fusion of the two sensors and the cross-entropy optimization on SICs are improving the CNN for prediction of sea ice. Improvements of predictions on high ice concentrations when adding MWR data make sense due to ambiguities in SAR data over homogeneous ice surfaces with open water under various wind conditions.

## VII. FUTURE WORK

Several limitations may still be addressed in future work. Currently, the noisy first subswath of the Sentinel-1 sensor may be left out under certain conditions, but future approaches could implement promising noise reduction techniques such as in [33] or using data processed with the ESA processor version 3.1 or larger (after June 2019). Further, experiments in the fusion of Sentinel-1 and AMSR2 may be expanded to account for non-linear relationships by adding additional layers at the end of the current model. An interesting additional study would be the importance of the different AMSR2 brightness temperatures in order to guide future missions in selecting the right frequencies for sea ice applications. There are numerous ways to do this though, and one could choose a leave-one-out or leave-one-in scheme and train CNNs for all possibilities. Another option is first converting MWR brightness temperatures to an estimated ice-concentration by a linear projection like in [19]. Further, one could calculate features on MWR data, such as channel ratios presented in [20], [34], to expand the possibilities of feeding the model with MWR derived information. However, it is unlikely to be feasible to do an exhaustive experimental search for an optimal selection of MWR data use.

Data augmentation techniques such as cropping overlapping window-samples from the large scenes, combined with larger CNN models, can possibly increase the performance in future experiments. More promising is it though, to incorporate a larger FOV in the network by increasing the dilation rate in the pooling layer of the CNN. Alternatively, one could consider trading a bit of resolution and changing the 40 m pixel spacing to, e.g. 80 m (which is closer to the Sentinel-1 SAR resolution) and thereby immediately double the FOV without increasing computational cost.

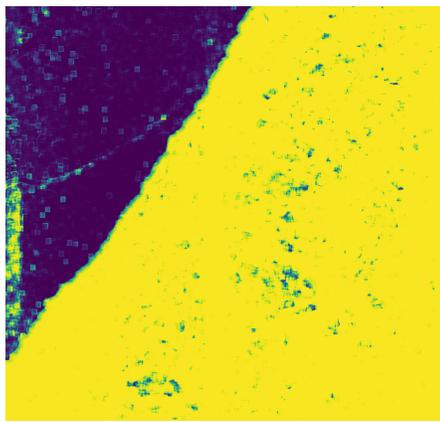
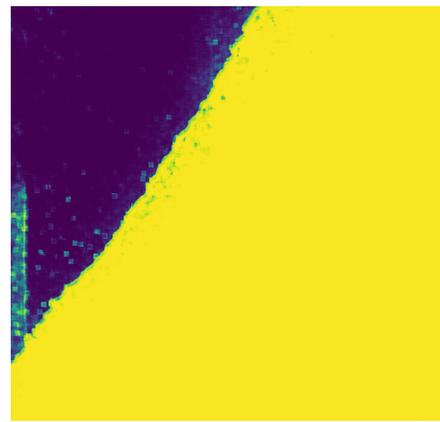
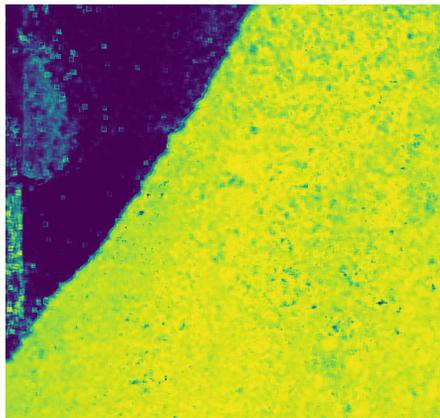
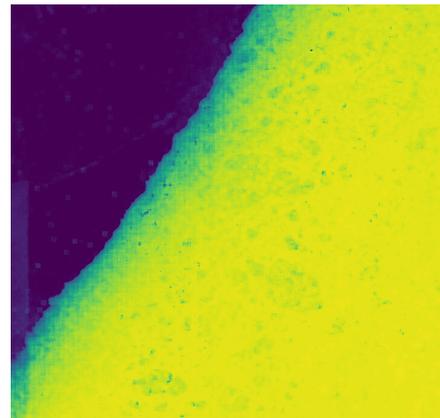
Finally experiments limited to specific regions (e.g. west Greenland) and/or seasons (separating Summer and Winter) could also constitute viable ways for improvements or investigation of the models performance under different ice conditions.

## ACKNOWLEDGEMENTS

The authors would like to thank the Innovation Fund Denmark for the financial support, Grant:7049-00008B, of the Danish research project ASIP (Automated Sea Ice Products) under which this work is carried out. ASIP has the aim to partially or fully automate the generation of sea ice information from Copernicus Sentinel-1 SAR imagery. We gratefully acknowledge the support of NVIDIA Corporation with the donation of GPUs used for this research, and we acknowledge the Japan Aerospace Exploration Agency (JAXA) for the use of the AMSR2 data.

## REFERENCES

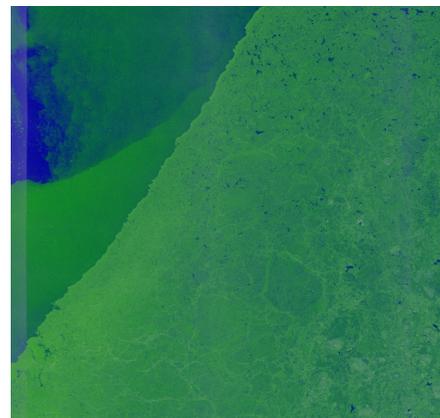
- [1] D. A. Clausi, "Comparison and fusion of co-occurrence, gabor and mrf texture features for classification of sar sea-ice imagery," *Atmosphere-Ocean*, vol. 39, no. 3, pp. 183–194, 2001.
- [2] L. Wang, K. A. Scott, and D. A. Clausi, "Sea ice concentration estimation during freeze-up from SAR imagery using a convolutional neural network," *Remote Sensing*, vol. 9, no. 5, p. 408, 2017.

(a) CNN Model predictions, S1,  $t = [T10\%]$ (b) CNN Model predictions, S1+AMSR2,  $t = [T10\%]$ (c) CNN Model predictions, S1,  $t = [SIC]$ (d) CNN Model predictions, S1+AMSR2,  $t = [SIC]$ 

(e) Colorbar representing the scale for prediction probabilities in 4 above figures.



(f) Icechart made by DMI ice analysts.



(g) RGB composite. Red: HV, Green: HH, Blue: HV/HH.

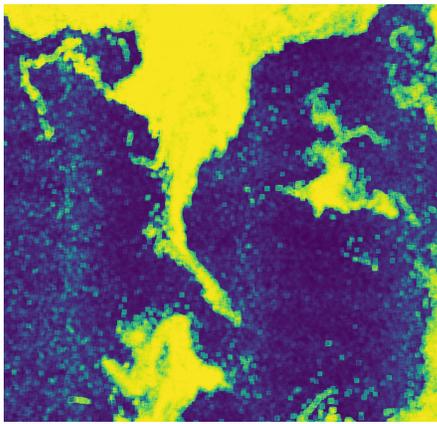
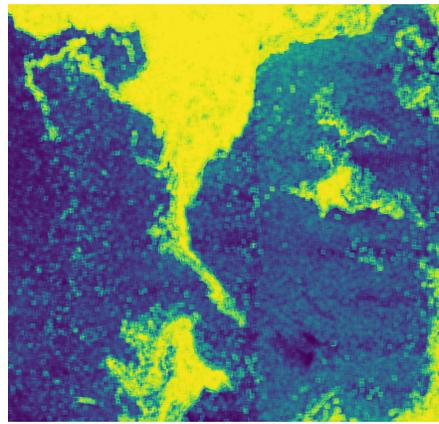
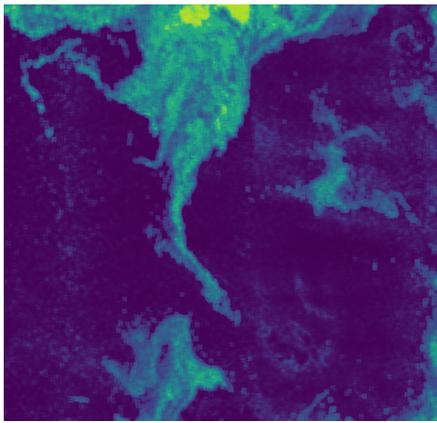
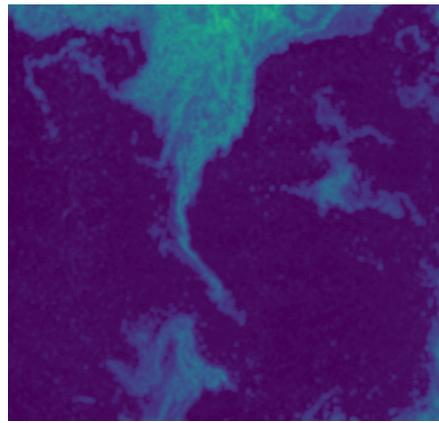
Fig. 5: Zoom of Sentinel-1 scene from 2015-06-17 (Product Unique ID: D45E) located in the sea between Greenland's west coast and Baffin Island. The size of the depicted area is 171 km x 180 km.

[3] D. Malmgren-Hansen, L. T. Pedersen, A. A. Nielsen, H. Skriver, R. Saldo, M. B. Kreiner, and J. Buus-Hinkler, "Asip sea ice dataset - version 1," March 2020. [Online]. Available: [https://data.dtu.dk/articles/ASIP\\_Sea\\_Ice\\_Dataset\\_-\\_version\\_1/11920416/1](https://data.dtu.dk/articles/ASIP_Sea_Ice_Dataset_-_version_1/11920416/1)

[4] World Meteorological Organization, "Instrument: Amsr-2," accessed, 2019/12/22. [Online]. Available: [https://www.wmo-sat.info/oscar/](https://www.wmo-sat.info/oscar/instruments/view/28)

[instruments/view/28](https://www.wmo-sat.info/oscar/instruments/view/28)

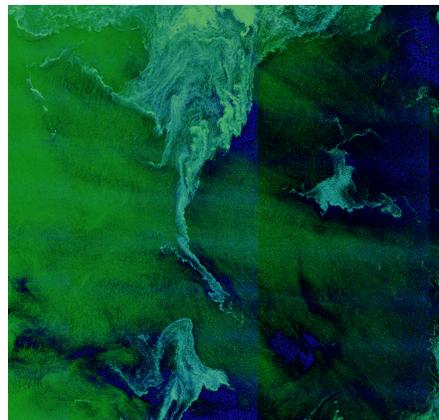
[5] J. Karvonen, J. Vainio, M. Marnela, P. Eriksson, and T. Niskanen, "A comparison between high-resolution eo-based and ice analyst-assigned sea ice concentrations," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 4, pp. 1799–1807, 2015.

(a) CNN Model predictions, S1,  $t = [T10\%]$ .(b) CNN Model predictions, S1+AMSR2,  $t = [T10\%]$ .(c) CNN Model predictions, S1,  $t = [SIC]$ .(d) CNN Model predictions, S1+AMSR2,  $t = [SIC]$ .

(e) Colorbar representing the scale for prediction probabilities in 4 above figures.



(f) Icechart made by DMI ice analysts.



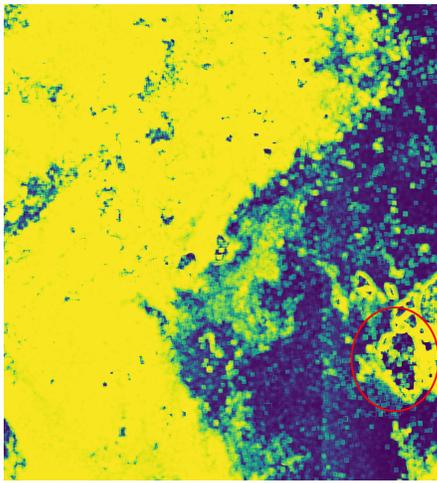
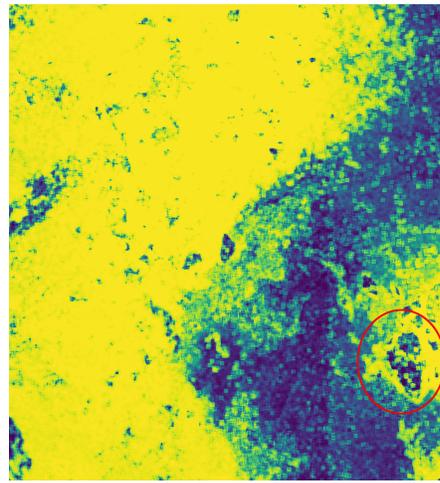
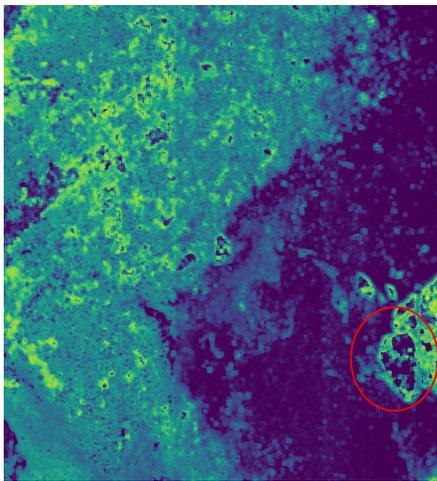
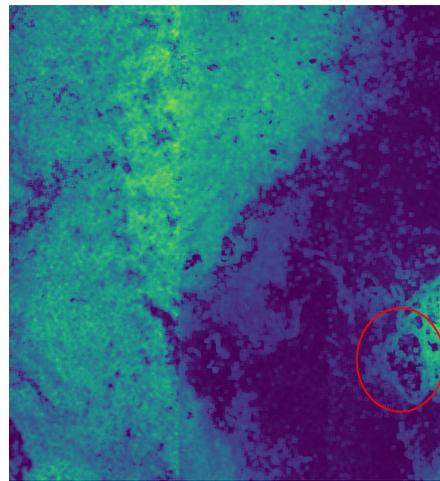
(g) RGB composite. Red: HV, Green: HH, Blue: HV/HH.

Fig. 6: The Sentinel-1 scene used in the above is from the 2017-08-22 (Product Unique ID: F1BD) located in the sea between the east coast of Greenland and the island Spitsbergen. The figures shows a zoom of the whole scene of 183 km x 192 km.

- [6] C. Pohl and J. L. Van Genderen, "Review article multisensor image fusion in remote sensing: concepts, methods and applications," *International journal of remote sensing*, vol. 19, no. 5, pp. 823–854, 1998.
- [7] J. Dong, D. Zhuang, Y. Huang, and J. Fu, "Advances in multi-sensor data fusion: Algorithms and applications," *Sensors*, vol. 9, no. 10, pp. 7771–7784, 2009.

- [8] H. Ghassemian, "A review of remote sensing image fusion methods," *Information Fusion*, vol. 32, pp. 75–89, 2016.
- [9] J. PRICE, "Combining panchromatic and multispectral imagery from dual resolution satellite instruments," *Remote Sensing of Environment*, vol. 21, no. 2, pp. 119–128, 1987.
- [10] A. Sarkar, A. Banerjee, N. Banerjee, S. Brahma, B. Kartikeyan,

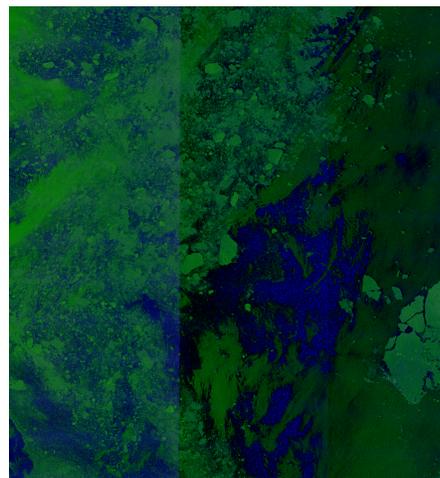
- M. Chakraborty, and K. L. Majumder, "Landcover classification in mrf context using dempster-shafer fusion for multisensor imagery," *IEEE transactions on image processing*, vol. 14, no. 5, pp. 634–645, 2005.
- [11] J. A. Benediktsson, J. R. Sveinsson, and P. H. Swain, "Hybrid consensus theoretic classification," *IEEE Transactions on geoscience and remote sensing*, vol. 35, no. 4, pp. 833–843, 1997.
- [12] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, "A review of semantic segmentation using deep neural networks," *International journal of multimedia information retrieval*, vol. 7, no. 2, pp. 87–93, 2018.
- [13] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.
- [14] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [15] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv preprint arXiv:1706.05587*, 2017.
- [16] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.
- [17] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian, "A real-time algorithm for signal analysis with the help of the wavelet transform," in *Wavelets*. Springer, 1990, pp. 286–297.
- [18] T. Lindeberg, "On the axiomatic foundations of linear scale-space," in *Gaussian scale-space theory*. Springer, 1997, pp. 75–97.
- [19] T. Lavergne, A. M. Sørensen, S. Kern, R. Tonboe, D. Notz, S. Aaboe, L. Bell, G. Dybkjær, S. Eastwood, C. Gabarro *et al.*, "Version 2 of the eumetsat osi saf and esa cci sea-ice concentration climate data records," *Cryosphere*, vol. 13, no. 1, pp. 49–78, 2019.
- [20] J. Karvonen, "Baltic sea ice concentration estimation using sentinel-1 SAR and AMSR2 microwave radiometer data," *IEEE Transactions on*
- Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 2871–2883, 2017.
- [21] C. M. Bishop, *Pattern recognition*, 1st ed., 2006, vol. 128.
- [22] N. Zakhvatkina, A. Korosov, S. Muckenhuber, S. Sandven, and M. Babiker, "Operational algorithm for ice–water classification on dual-polarized radarsat-2 images," *The Cryosphere*, vol. 11, pp. 33–46.
- [23] S. Leigh, Z. Wang, and D. A. Clausi, "Automated ice–water classification using dual polarization sar satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5529–5539, 2013.
- [24] L. Wang, K. A. Scott, and D. A. Clausi, "Improved sea ice concentration estimation through fusing classified sar imagery and amsr-e data," *Canadian Journal of Remote Sensing*, vol. 42, no. 1, pp. 41–52, 2016.
- [25] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.
- [26] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 8–36, 2017.
- [27] A. Lagrange, B. Le Saux, A. Beaupere, A. Boulch, A. Chan-Hon-Tong, S. Herbin, H. Randrianarivo, and M. Ferecatu, "Benchmarking classification of earth-observation data: From learning explicit features to convolutional networks," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2015, pp. 4173–4176.
- [28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [29] D. Malmgren-Hansen, "Asip project code," <https://github.com/damaha/asip-v1.git>, 2019.
- [30] K. Harnvig, "Asip user survey report," [https://asip.dk/onetwebmedia/ASIP\\_WP10\\_Fase-1-rapport\\_08.04.2020.pdf](https://asip.dk/onetwebmedia/ASIP_WP10_Fase-1-rapport_08.04.2020.pdf), 2020.
- [31] C. Sun, A. Shrivastava, S. Singh, and A. Gupta, "Revisiting unreasonable effectiveness of data in deep learning era," in *Computer Vision (ICCV), 2017 IEEE International Conference on*. IEEE, 2017, pp. 843–852.
- [32] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *arXiv preprint arXiv:1503.02531*, 2015.
- [33] J.-W. Park, J.-S. Won, A. A. Korosov, M. Babiker, and N. Miranda, "Textural noise correction for sentinel-1 topsar cross-polarization channel images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 6, pp. 4040–4049, 2019.

(a) CNN Model predictions, S1,  $t = [T10\%]$ .(b) CNN Model predictions, S1+AMSR2,  $t = [T10\%]$ .(c) CNN Model predictions, S1,  $t = [SIC]$ .(d) CNN Model predictions, S1+AMSR2,  $t = [SIC]$ .

(e) Colorbar representing the scale for prediction probabilities in 4 above figures.



(f) Icechart made by DMI ice analysts.



(g) RGB composite. Red: HV, Green: HH, Blue: HV/HH.

Fig. 7: The Sentinel-1 scene used in the above is from the 2017-08-11 (Product Unique ID: EF7C) located off the east coast of Greenland near Shannon island. The figures shows a zoom of the whole scene of 244 km x 224 km. A red circle in the predictions mark an area where the networks fail to classify ice properly.

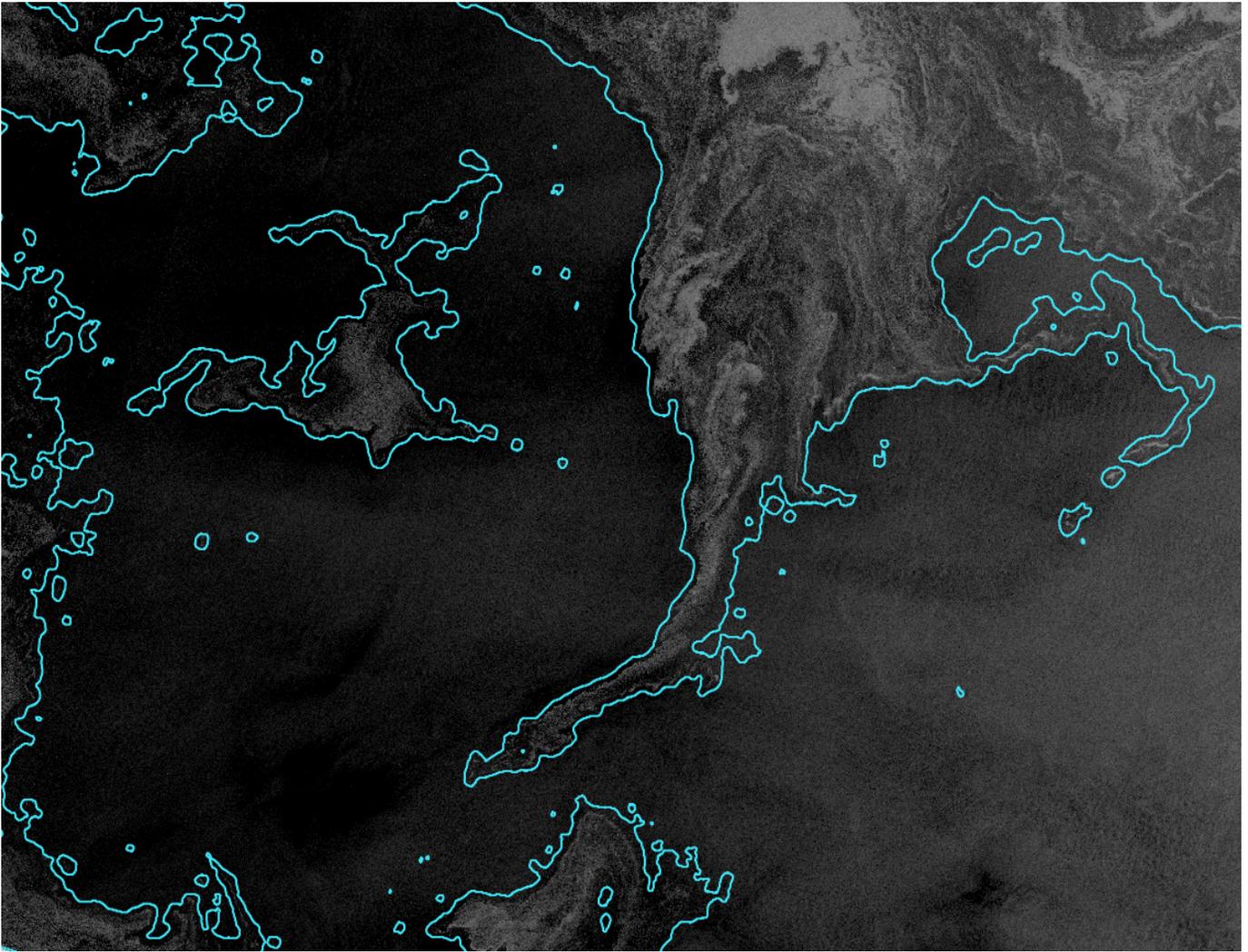


Fig. 8: Polygon outline extracted from CNN predictions (model trained with SICs including both S1 and AMSR2) overlaid on Sentinel-1 scene from the 2017-08-22 (PID:F1BD), HH channel. A 3x3 kernel Gaussian smoothing has been applied to the predictions prior to extraction of the outline. The image is geographically aligned in QGIS.