U-net with ResNet-34 backbone for dual-polarized C-band baltic sea-ice SAR segmentation

Juha Karvonen

Finnish Meteorological Institute (FMI), Helsinki, Finland

Abstract

In this study, the U-net with ResNet-34, i.e. a residual neural network with 34 layers, backbone semantic segmentation network is applied to C-band sea-ice SAR imagery over the Baltic Sea. Sentinel-1 Extra Wide Swath mode HH/HV-polarized SAR data acquired during the winter season 2018–2019, and corresponding segments derived from the daily Baltic Sea ice charts were used for training the segmentation algorithm. C-band SAR image mosaics of the winter season 2020–2021 were then used to evaluate the segmentation. The major objective was to study the suitability of semantic segmentation of SAR imagery for automated SAR segmentation and also to find a potential replacement for the outdated iterated conditional modes (ICM) algorithm currently in operational use. The results compared to the daily Baltic Sea ice charts and the operational ICM segmentation and visual interpretation were encouraging from the operational point of view. Open water areas were located very well and the oversegmentation produced by ICM was significantly reduced. The correspondence between the ice chart polygons and the segmentation results was also reasonably good. Based on the results, the studied method is a potential candidate to replace the operational ICM SAR segmentation used in the Copernicus Marine Service automated sea-ice products at Finnish Meteorological Institute.

1. Introduction

Sea ice is a crucial component of the global climate system, and its extent and distribution play an important role in the exchange of energy, momentum and mass between the ocean and the atmosphere. Sea-ice information is also important and necessary for navigation and other offshore activities in sea ice. Synthetic Aperture Radar (SAR) is a powerful tool for monitoring sea ice due to its ability to operate in all weather conditions and its sensitivity to the different physical properties of the ice. SAR images can be used to derive information on the extent, concentration and type of sea ice, for example. However, the interpretation of SAR images is often challenging due to the noisy nature of SAR imagery, complexity of the ice cover and the presence of other types of targets, such as open water, waves in open water, icebergs and vessels in SAR scenes. Therefore, image segmentation techniques are important for the analysis of SAR images of sea ice. Finnish Meteorological Institute (FMI) is involved in the Copernicus Marine Service (CMS) sea-ice thematic assembly center (SITAC) as the Baltic Sea processing unit. CMS is a European Commission funded operational service coordinated by Mercator Ocean International (MOI), located in Toulouse, France. The role of FMI in CMS SITAC is to provide timely sea-ice information over the Baltic Sea, based on SAR data and other available information on sea ice. The FMI sea-ice concentration (SIC) (Karvonen, 2017) and sea-ice thickness (SIT) (Karvonen and others, 2003) CMS SITAC products provide operational segmentwise near-real-time (NRT) estimates for SIC and SIT. SAR segmentation is an essential step of making these products.

Image segmentation is the process of dividing an image into meaningful uniform regions, each of which corresponds to a different object or a uniform part of the scene. In the context of sea-ice SAR images, segmentation can be used to separate different types of ice, such as first-year ice, multi-year ice and icebergs. Several segmentation techniques have been applied to sea-ice SAR images, ranging from classical approaches, such as thresholding and clustering, to more advanced methods, such as neural networks and deep learning. In general, segmentation just divides an image into separate uniform areas. Semantic segmentation, in addition to this, assigns a class or category, such as sea ice and open water in the case of sea-ice imagery, to each segment.

Thresholding is a simple yet efficient technique for segmenting images. It involves selecting one or more threshold values that separate the pixels belonging to different classes based on their intensity or other local image features. For example, the threshold can be set to separate ice pixels from open water pixels based on their SAR backscatter values. Several studies have applied thresholding to sea-ice SAR images using different threshold selection approaches, a short overview of thresholding techniques for SAR segmentation is given, e.g. in Al Bayati and El Zaart (2013). The drawback of direct simple thresholding is that it is sensitive to speckle and thermal noise present in SAR images. The different backscattering due to incidence angle variation affects the thresholding, also depending on the scattering surface (ice or open water) structure (Makynen and Karvonen, 2017). Efficient segmentation based on thresholding would

require filtering of both speckle and thermal noise before applying. In Lee and Jurkevich (1988) thresholding based on histograms applied to SAR images with different numbers of looks was studied, also with speckle filtering applied before segmentation. The study was performed using synthetic SAR data and incidence angle was not taken into account. The results apply only for a narrow incidence angle ranges and indicate that thresholding based on grayscale histogram is applicable for SAR segmentation. However, in the case of a wide incidence angle range, the histogram calculation and segmentation should be performed for multiple incidence angle sub-ranges. For sea ice it is possible to apply incidence angle dependence correction (Makynen and Karvonen, 2017) before thresholding but for open water correction it is not possible because backscattering from open water is dependent on the instantaneous local wave spectrum and it is in practice unknown.

Clustering is another commonly used traditional technique for segmenting SAR images. Clustering algorithms group pixels based on their similarity in terms of intensity, texture or other local features. The K-means algorithm (MacQueen, 1967) is a simple and popular clustering technique that can be applied to SAR imagery also. K-means has been applied to sea-ice SAR, e.g. in Yu and others (2013); Ren and others (2015); Zhang and Skjetne (2015). An obvious problem with K-means is that it requires the number of clusters, K, in advance as its input. K-means implicitly assumes that clusters are roughly spherical (isotropic, uniform in all directions) and equally sized, but this may not exactly be true, depending on the data points. To reduce the effect of this property of K-means, multiple clusters can be used for a class. Also to overcome this restriction of K-means, K-means can be modified to take a mixture of Gaussian classes (with different covariances) into account by using the Mahalanobis distance instead of the Euclidean distance (Brown and others, 2022). Also unsupervised K-means versions to overcome this requirement have been proposed, e.g. in Sinaga and Yang (2020). Many more advanced clustering techniques have been developed and applied to different types of image data, e.g. Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH) (Zhang and others, 2006) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester and others, 2006). Clustering based on the meanshift (MS) algorithm (Cheng, 1995) is used at FMI as an operational SAR processing step to provide an initial clustering and a suitable number of clusters before applying a contextual segmentation.

Contextual segmentation methods take into account the neighborhood (context) of each pixel and perform better for noisy images when compared to straightforward pixel-wise thresholding or clustering approaches. Examples of contextual methods are the Iterated Conditional Modes (ICM) algorithm Besag (1986), Markov Random Field (MRF) based methods, applied to sea-ice SAR, e.g. in Deng and Clausi (2005); Maillard and others (2006) and Pulse-Coupled Neural Network (PCNN), applied to sea-ice SAR, e.g. in Karvonen (2004). The current operational SAR segmentation applied at FMI uses the ICM algorithm (Karvonen, 2017). ICM was originally selected at FMI because the segmentation result was at an acceptable level for FMI operational purposes and ICM performed significantly faster than considered MRF models that also produced acceptable segmentation results.

There exist a wide variety of different image segmentation techniques, also applicable to SAR data. An overview of general image segmentation techniques with references can be found, e.g. in Jeevitha and others (2020) and Yu and others (2023). Many different kinds of segmentation methods have also been applied to SAR imagery. Just to mention a few examples of different methods applied to SAR imagery segmentation: such examples are edge-based segmentation (Oliver, 1994), watershed

segmentation (Li and others, 1989), and segmentation based on active contours (Ayed and others, 2004).

More advanced and complex segmentation techniques, such as neural networks and deep learning, have also been applied to sea-ice SAR images during the recent years. These techniques are based on the use of artificial intelligence algorithms that learn to identify different features of the images by training with a large dataset of labeled images. For example, convolutional neural networks to segment SAR images and specifically sea-ice SAR images have been applied in Malmgren-Hansen and Nobel-Jorgensen (2015); Dowden and others (2020); Boulze and others (2020); Zhang and Chen (2022); Wang and Li (2021); Karvonen (2021); Wan and others (2023). Typically these deep learning approaches provide a semantic segmentation, i.e. also include a categorization of the segments. A general overview of deep learning applied to SAR has been published in 2021 (Zhu and others, 2021). The U-net (Ronneberger and others, 2015) and its variants have gained popularity in image segmentation during the recent years, and it has also been applied to sea-ice SAR imagery, e.g. in Stokholm and others (2022) and Huang and others (2021).

Transformer networks were originally proposed by Vaswani and others (2017). The idea of a transformer network, or shortly just transformer, is based on parallel multi-head attention. Attention mechanism is a kind of neural network layer that can be included in deep learning models. Attention allows the model to focus on specific parts of the input by assigning different weights to different parts of the input. This weighting is typically based on the relevance of each part of the input to a specific task. The attention mechanism allows to include large-scale context into a neural network in an efficient way. Transformers were originally developed for linguistic tasks. For image data, vision transformers were later developed (Dosovitskiy and others, 2021). There also exits a recently published survey on using transformers in remote sensing (Aleissaee and others, 2021). Ren and others (2023) applied dual-attention mechanism incorporated into the U-net architecture to SAR data. Sea-ice SAR segmentation based on transformer networks has been applied, e.g. in Ristea and others (2023); Li and others (2023).

In conclusion, image segmentation is a significant step for the analysis of SAR images of sea ice and there exists a wide variety of different segmentation methods applicable to SAR data. Actually, manual ice charting typically uses a similar approach as SAR segmentation: in the manual ice analysis the area of interest is first divided into areas representing uniform ice regions presented as polygons. After this step, attributes describing the ice within these areas are assigned to these defined regions. Applying automated SAR segmentation techniques also provides a good basis for efficient segment-wise analysis of sea ice either by machine or by human ice analysts.

The aim of this study was to find a suitable algorithm for sea-ice SAR segmentation to be used as the segmentation step to produce automated SAR based sea-ice information products and be applicable in the operational CMS SITAC product processing. The SAR segmentation step also roughly corresponds to the defining and drawing of the polygons in the daily manual Baltic Sea ice charting process. The ice charts are presented in a vector graphics format (JCOMM, 2014) and the uniform areas are presented there as polygons, defined by sets of coordinate (x, y) pairs. After the segmentation (or polygonalization) phase there exists a set of regions each representing approximately uniform ice conditions. The attributes describing the segment content will then be assigned to the segments in a later phase of the automated processing chain. The intended context of the SAR segmentation at FMI as part of the automated sea-ice information production chain is demonstrated in Fig. 1.
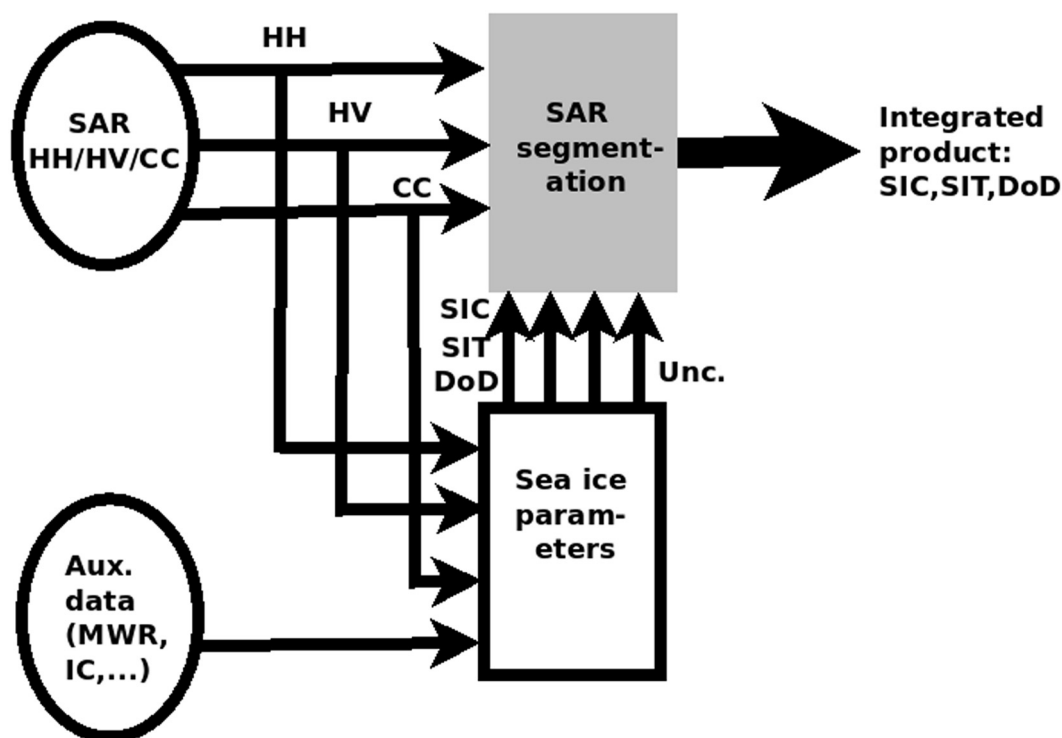
Figure 1. Simplified diagram of the context in which SAR segmentation (gray box) will be used at FMI operational SAR production chain. The SAR data are used for segmentation and for estimating sea-ice parameters together with some complementary auxiliary data, e.g. from microwave radiometer. IC in the figure refers to the digitized ice chart, CC to the HH/HV SAR channel cross-correlation. Necessary sea-ice parameters are, e.g. ice concentration (SIC), sea-ice thickness (SIT) and its distribution, degree of ice deformation (DoD). Uncertainties (unc.) of the parameter estimation are also important. These values are then assigned to each single segment as segment medians or segmentwise parameter value distributions and the final output will be an integrated product with multiple layers corresponding to estimated sea-ice parameters.

ResNet (He and others, 2016), or residual neural network, is actually a family of residual neural networks with different architectures and numbers of layers, indicated by their names, e.g. ResNet-34, ResNet-50, ResNet-101, ResNet-152, getting computationally more complex with an increasing number of layers. It was found that the U-net with ResNet-34 backbone (denoted in this paper by U-net/ResNet-34), applied e.g. in Ren and others (2023), provided visually reasonable segmentation results in preliminary tests when applied to the FMI dual-polarized C-band sea-ice SAR imagery. ResNet with more convolution layers would probably have provided slightly more accurate results but ResNet-34 provides a good compromise between computational complexity and segmentation performance. Based on the preliminary small-scale tests, U-net/ResNet-34 was selected for further tests with the FMI operational SAR data. The computational complexity of the selected method is still suitable for a standard desktop computer used for computations in this particular study. This study is based on tests performed with Sentinel-1 (training data) and combined Sentinel-1 and Radarsat-2 imagery (test data).

One obvious advantage of applying a segmentation and then using segment-wise sea-ice information is the ability to compress the information grid efficiently, e.g. compared to SAR imagery with poorly compressive speckle noise and many spatially and temporally local features caused by the instantaneous rapidly changing wave conditions over open water. A good compression is important, e.g. when delivering ice information to ships via a satellite connection. Another advantage is that automated methods perform systematically. Ice charts made by ice analysts are dependent on the varying interpretations of different ice analysts. Large variation between ice analyses made by different ice analysts have been demonstrated and reported, e.g. in Karvonen and others (2015).

2. Study area and datasets

The study area, Baltic Sea (see Fig. 2), is a semi-enclosed brackish water basin with a seasonal ice cover located in northern Europe, approximately between latitudes 53 N and 65 N and longitudes 9E and 30E. The area of Baltic Sea is 422 000 km^2, and average annual maximum ice cover extent is 170 000 km^2. Many of the harbors in the Baltic Sea are ice-surrounded every winter and precise and timely ice information is necessary for navigation in the ice-covered areas. The winter ship traffic is maintained with the aid of ice breakers. A typical Baltic Sea ice season lasts from November to December until late May in the northern parts (Gulf of Bothnia). The thermodynamically grown ice in the fast ice zone is at maximum around 1 m thick, on average 72 cm (Ronkainen, 2013), in the northern Gulf of Bothnia. In deformed ice areas, ice thickness can be several meters, in ice ridges even 25 m (Kankaanpaa, 1997; Granskog and others, 2006). The maximum ice extent in Baltic Sea is typically reached in February–March.

The digitized FMI ice charts have been used in this study to derive training data for the proposed SAR segmentation algorithms and also as reference data for evaluation of the segmentation algorithm. In ice charts, ice parameters are estimated by ice analysts for polygons they draw on a map base according to their interpretation of the ice conditions. Each polygon represents an ice type or multiple ice types which can uniquely be described in terms of the ice charting guidelines provided by the World Meteorological Organization (WMO) (JCOMM, 2014). SIC is also assigned to each of these polygons. In ice charting based on the WMO guidelines, one polygon can involve more than one ice type and thus also multiple concentrations of these multiple ice types. In ice charts, polygon-wise sea-ice information is
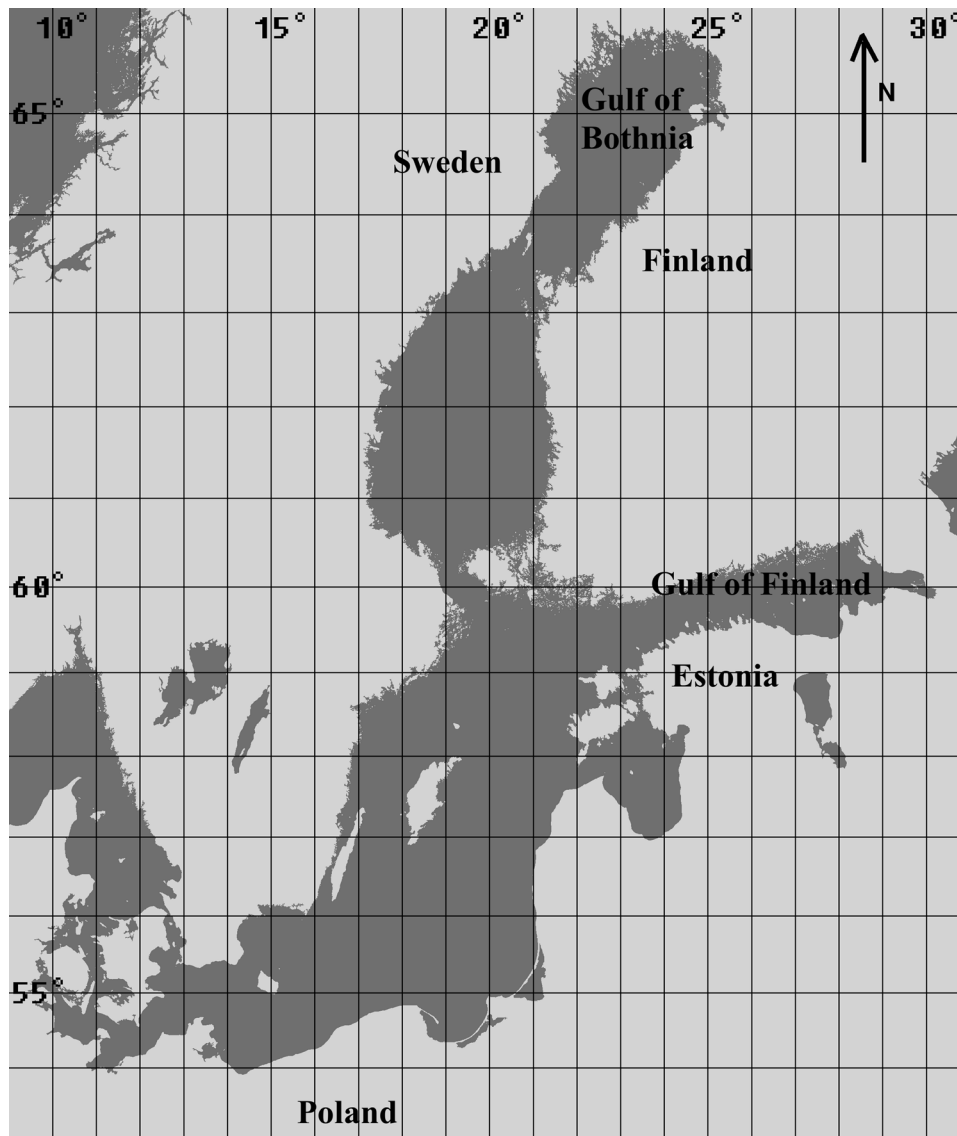
Figure 2. Study area, the Baltic Sea.

typically indicated by the egg code assigned to each polygon (Canadian Ice Service, 2005). The FMI ice charts over the Baltic Sea are made daily by ice analysts during the winter period. The input data for making the FMI ice charts are satellite data from multiple instruments, including SAR (C band: Sentinel-1, RADARSAT-2, Radarsat Constellation Mission, and X band: COSMO-SkyMed, TerraSAR-X, PAZ) and optical/infrared data (MODIS, VIIRS), and observation data from coastal observers, observations from the Finnish and Swedish ice breakers and the FMI operational sea-ice forecast model results. The most important data source is the Sentinel-1 C-band SAR images provided in near-real-time by the European Space Agency (ESA). In the FMI ice charts, the ice analyst first locates the areas with homogeneous ice conditions which are presented by polygons. Then, for each ice chart polygon five attributes (ice concentration, ice minimum thickness, ice average thickness, ice maximum thickness, degree of ice deformation) are assigned by the ice analyst. In the FMI Baltic Sea ice charts, the given SIC is the total SIC of the polygon given in percents, not in tenths as in the ice charts according to the standard WMO guidelines (provided in ice chart egg code diagrams). In the FMI Baltic Sea ice charts, neither partial SIC nor stage of ice development is given. Instead the five attributes listed above are assigned to each polygon; for open water areas

(polygons), the sea surface temperature (SST) attribute is assigned. The SIT values are level ice thicknesses, i.e. values for thermodynamically grown ice, possibly drifted from their original locations. Sea-ice deformation (SID) is given as a five-stage scale in which one represents smooth level ice and five highly deformed ice. The thickness of deformed ice can roughly be estimated by multiplying the level ice thickness by the degree of deformation. For example, rafted ice, with two overlying level ice layers would then correspond to SID value of two. The ice classes used in this study were derived from the SIC, SIT and SID fields of the ice charts; the derivation is described in more detail in Section 3. In practice, there appeared ten major sea-ice classes in the training data covering the whole ice season 2018–2019. These classes are described in Table 1. The numbers of the classes in the table and used later in this paper are just order numbers resulting from the quantization of three independent digitized ice chart parameters (sea-ice concentration, sea-ice thickness, degree of deformation) available for each ice chart polygon. These class numbers do not have any other specific meaning, except that in general the class number increases as a function of decreasing ice navigability. It should be noted that these classes are the ice classes used for training the U-net/ResNet-34 algorithm. Other possible classes based on this quantization were

Table 1. Sea-ice classes

Class number	properties
1	Open water
8	Thin ice (SIT 1–20 cm) , SIC 20–40%
9	Thin ice, SIC 50–60%
10	Thin ice, SIC 70–90%
11	Thin ice, SIC 100%
18	SIT 20–40 cm, SIC 90–100%
24	SIT 40–60 cm, SIC 90–100%
26	SIT 20–40 cm, SIC 80–90%, deformed
28	SIT 20–40 cm, SIC 90–100%, deformed
32	SIT 40–60 cm, SIC 90–100%, deformed

The class numbers correspond to the class numbers used throughout this study.

very rare or did not exist at all in the training data and they were excluded from the training and classification.

The SAR training data used in this study were Sentinel-1 extra wide swath (EW) Ground Range Detected Medium resolution (GRDM) dual polarization mode level 1B (L1B) data (Bourbigot and others, 2016). The two channels of the images represent the HH and HV polarization combinations, where the first letter indicates the transmitted polarization and the second letter received polarization and H is horizontal polarization and V is vertical polarization. The swath width of the used acquisition mode was about 400 km. The training SAR data of this study consisted of 234 Sentinel-1 images. This number of imagery experimentally proved to be large enough for training the network and also suitable for the computing resources in use, i.e. a desktop computer without any specific hardware for neural computing. With a dataset consisting of about half of this training dataset used in this study, the segmentation test results were still too poor for any practical use. The number of monthly scenes in the training imagery of the 2028–2019 winter season training imagery is shown in Figure 3. The final training data were cropped to 256×256 pixel windows and they were used as inputs to the U-net/ResNet-34. The training dataset was augmented by applying vertical and horizontal flip and rotation in 90 degree steps, i.e. four rotations (0, 90, 180 and 270 degrees) of each 256×256 pixel training window were used. To test the segmentation an independent dataset consisting of daily SAR mosaics of the winter 2020–2021 was used. Actually, during this test winter, the ice formation in Baltic Sea started quite late and significant amounts of sea ice in the Baltic appeared only in January 2021, so mosaics over a period from January to May 2021 were used
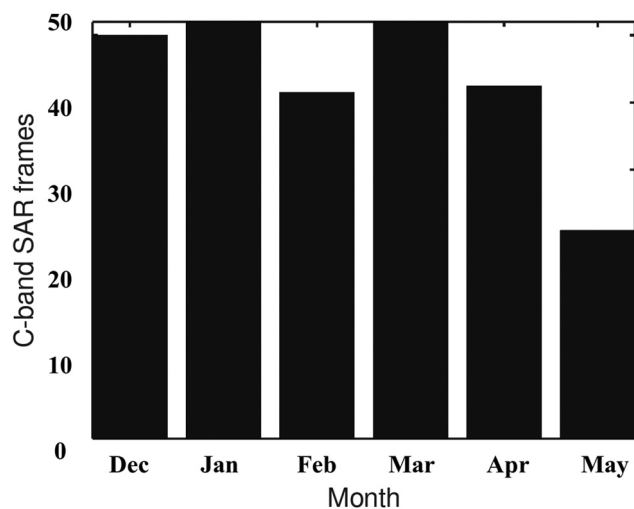

Figure 3. Monthly distribution of the amount of winter 2018–2019 SAR images used in this study.

in the tests. The daily SAR mosaics were made by always overlaying the most recent SAR image and thus the most recent available SAR measurement was available at each mosaic grid point. Separate mosaics were generated for HH and HV channels and their resolution was 500 m. Both Radarsat-2 ScanSAR HH/HV polarized (MDA, 2018) and Sentinel-1 EW GRDM HH/HV mode L1B images were used to generate the SAR mosaics.

The training images were selected randomly from all the winter season 2018–2019 data of about 650 Sentinel-1 EW GRDM mode HH/HV images and they represent both cold conditions and ice melt conditions. The reference dataset for all the experiments were the daily digitized FMI ice charts of the same day with the SAR acquisitions and their polygons originally assigned to 32 classes based on the sea-ice properties provided by the ice analysts. The FMI ice charts were used to generate the training datasets and as reference data in the comparisons evaluating the performance of the segmentation algorithm.

The SAR data were calibrated, georectified into Mercator projection with WGS84 datum and 61 degrees 40 min reference (correct scale) latitude, the logarithmic SAR backscattering coefficient values, denoted by σ^0, were quantized to eight bits per pixel (8 bpp), such that for the HH channel σ^0 of -30 dB or less corresponds to pixel value of one and 0 dB to the pixel value of 255. For the HV band, mostly with a lower σ^0, the corresponding values were -40 and 0 dB. The pixel value zero was reserved for background (no data and land mask). This quantization has proved to be a good solution for sea-ice SAR classification and sea-ice parameter estimation from SAR imagery and been in use at FMI for a long time. According to tests made at FMI in 2016 there was in practice no difference between this 8 bpp presentation and 16 bpp presentation of the images in sea-ice parameter estimation and the 8 bpp presentation was selected, e.g. for the SIC estimation (Karvonen, 2017). Before the quantization, a linear incidence angle correction based on the slopes provided in Makynen and Karvonen (2017) was performed. The land masking was performed based on a land mask derived from the Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG) coastline dataset (Wessel and Smith, 1996) applied to the georectified SAR images. Based on earlier experience, this quantization preserves the SAR texture well and with sufficient accuracy for automated classification. And with the channel-wise quantization ranges defined as above there will not appear large areas of pixels saturated to the upper (0 dB) or lower boundaries ($-30/-40$ dB). Similar quantization scheme applied to dual-polarized (HH/HV) C-band SAR data has earlier been used for example in Karvonen (2015) (for Radarsat-2 data) and in Karvonen (2017) (Sentinel-1) in the context of SAR texture-based SIC estimation. The 8 bpp data were then down-sampled to the resolution of 500 m. Also the HH/HV channel cross-correlation (CC) in the same 500 m resolution was computed. In this study, CC was computed using the quantized and downsampled (500 m resolution) SAR imagery. CC was used as a third channel of the images used in this study. The 8 bpp HH and HV channels and the HH/HV CC, computed in a round-shaped windows with a radius of three pixels and scaled from the range of [0,1] to [1,255] and rounded to the nearest integer, are the three image channels used as inputs to the segmentation. The image channels are similar as in Karvonen (2021).

3. Methodology

The U-net (Ronneberger and others, 2015) is a convolutional neural network based on the fully convolutional network (Shelhamer and others, 2017). In the U-net, the usual contracting network layers are supplemented by successive layers where pooling is replaced by upsampling. These layers increase the resolution

of the output. A successive convolutional layer is then able to learn a precise output based on its input. The U-net can be presented in the shape of letter U consisting of decoder and encoder blocks that are connected via so-called bridges or skip connections. A simplified schematic diagram of the U-net is presented in Figure 4.

In this study, the U-net with the convolutional backbone of a ResNet-34 (He and others, 2016) is applied. The term backbone refers to the feature extraction network processing the input data into a feature presentation (encoder). The structure of the encoder then determines the basic structure of the decoder part of the network. The idea of this combination is to integrate the use of a proven image classification architecture into the U-net. ResNet-34 is a convolutional neural network architecture proposed by Microsoft Research Asia in 2016. The architecture is based on the idea of residual learning, which allows training of much deeper neural networks with improved accuracy, compared to networks not utilizing residual learning. As indicated by its name, ResNet-34 has totally 34 convolution layers. In ResNet-34, the input image is first passed through a convolutional layer followed by a batch normalization layer and a rectified linear unit (ReLU) activation function. This is followed by a series of residual blocks, each of which consists of two convolutional layers, each followed by batch normalization and ReLU activation. Resnet-34 consists of five convolution (resolution) layers with multiple convolution filters at each layer. The output size after each convolution layer is downsampled by two in the two image window dimensions. The U-net with Resnet-34 backbone has an U-net skip connection after each ResNet-34 convolution layer. These layers form the encoder part of the network. Each block of the decoder part starts with an upsampling layer to increase the spatial resolution. The feature maps provided by the upsampling are then concatenated with the feature maps of the corresponding layer from the encoder side through the U-net skip connections. Following the concatenation, two convolutional layers. are applied to refine the feature maps. The two consecutive convolutional layers are followed by batch normalization and ReLU activation steps. The structure for the U-net with the ResNet-50 backbone has been described in Manos and others (2022). This structure is similar to U-net/RssNet-34 applied here, except for the number of applied convolution kernels.

The ResNet-34 residual blocks also contain a shortcut connection that allows the input to bypass the convolutional layers and be added to the output of the residual block. This shortcut connection helps to alleviate the so-called vanishing gradient problem
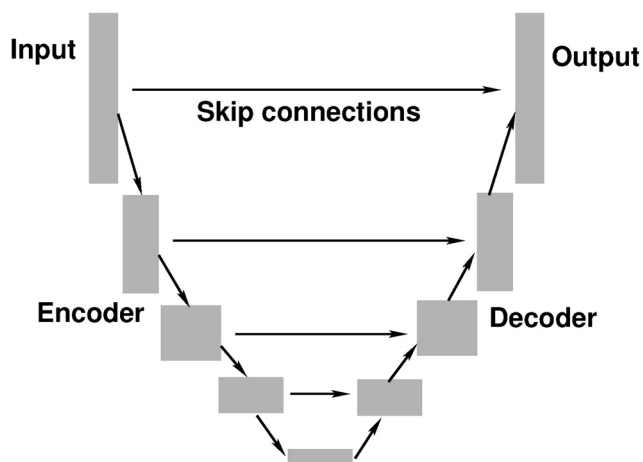
causing degradation of the learning process when increasing the number of layers of the network. Using residual blocks in the neural network enables training of deeper networks. In gradient-based learning algorithms, gradients are used to learn the weights of a neural network. It works like a chain reaction as the gradients closer to the output layers are multiplied with the gradients of the layers closer to the input layers. These gradients are used to update the weights of the neural network. If the gradients are small, the multiplication of these gradients will become so small that it will be close to zero. This results in the model being unable to learn, and its behavior becomes unstable. This problem is called the vanishing gradient problem (Hochreiter, 1991).

The input to a residual (or skip) ReLU layer from the previous layer is here denoted by x and the layer output by $F(x)$, and the expected output of the layer by $H(x)$. In a residual ReLU block sum of x and $F(x)$ is fed to the next layer. This means that $H(x) = F(x) + x$ and $F(x) = H(x) - x = R(x)$. $R(x)$ is the residual. This indicates that the residual layers are actually trying to learn the residual $R(x)$. A simplified residual layer structure is shown in Figure 5.

The classes were defined based on the FMI daily ice chart SIC (C_{ice}), ice thickness (H_{ice}) and deformation (D_{ice}) by quantizing these three independent quantities as follows.

$$
\begin{aligned}
C_1 &= (C_{ice} + 4)/10 + 1 \\
C_2 &= Min(H_{ice}/20 + 1, 5) \\
C_3 &= H(D_{ice} - 2 - \epsilon) + 1
\end{aligned}
\tag{1}
$$

$H(x) = (\text{sign}(x) + 1)/2$ is the Heaviside function, i.e. one for positive x and zero for negative x, ϵ is a small positive value, here just less than one. The theoretical ranges of the sub-categories C_1, C_2 and C_3 are (1,11), (1,5) and (1,2), respectively; this would result to a maximum number of 110 possible classes, but in practice for a typical Baltic Sea ice winter much less classes according to the ice charts appear. In practice, there appeared ten of these classes in the training data. The number of other classes was neglectably small or they did not appear at all.

The processing of the training input data is already explained in the previous section and the training phase is shown in Figure 6. The images are cropped into 256×256 pixel windows,

Figure 4. A schematic diagram of the U-net structure. Essential for the U-net are the skip connections between different encoder (on the left) and decoder (on the right) levels representing different resolutions decreasing downwards.
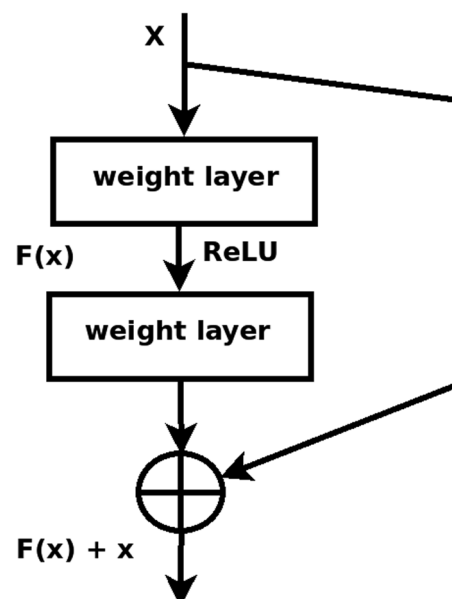

Figure 5. A residual ReLU block. Residual neural networks use this kind of block to allow robust learning of deeper networks. Weight layer in the figure indicates a convolution layer.
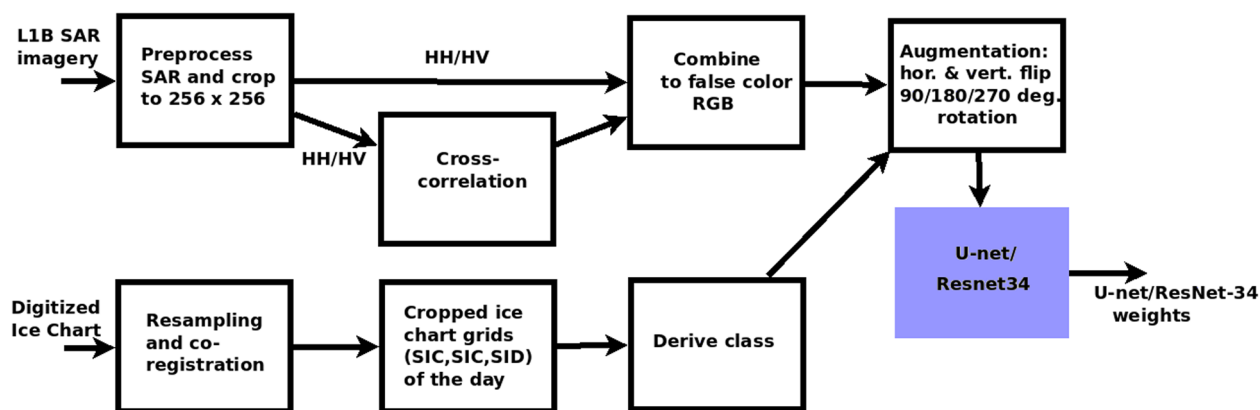
Figure 6. Diagram of the training phase. L1B SAR data are first preprocessed, including calibration, thermal noise filtering, georectification to Mercator projection and resampling to 500 m resolution.

the corresponding ice chart grids are cropped the same way (same location). For the cropped SAR images HH/HV cross-correlation (CC) channel is computed and HH, HV and CC are combined to a false color RGB image (R = HH, G = HV, B = CC). For the ice chart 256 × 256 pixel blocks the classes are derived based on the ice chart quantities assigned to the ice chart polygons and the classes of Table 1 defined by the quantification of Eqn (1). Data augmentation (flips and rotations) is applied to each 256 × 256 RGB image block and ice chart class block. In the estimation phase, a 256 × 256 sliding window with a step of 200 pixels is used in segmentation and at the overlapping boundaries 28 pixels of both the overlapping 256 × 256 windows are not included in the segmentation. Splitting with an overlap is done to avoid possible artifacts near the cropped image boundaries. Some undesired artifacts (fake segments) seemed to appear near the window boundaries. Using overlapping windows and ignoring the window boundary areas seemed to remove these artifacts in the training and validation datasets. This also worked for the independent test dataset consisting of season 2020–2021 SAR mosaics. The segmentation block diagram in Figure 7 contains similar processing steps as in the training phase: CC channel is computed and HH, HV and CC are combined to an RGB image, cropped to 256 × 256 windows and fed to the U-net/ResNet-34, and finally the cropped images are combined to a full size image again.

The software library used to implement the algorithm was the python3 segmentation_models library (Iakubovskii, 2019), and its U-net with ResNet-34 backbone module. The pre-trained weights available in the package were not applied in this study because

SAR imagery has different properties than the optical imagery used for the pre-training. Several loss functions were tested. The results with different loss functions were rather similar. For example, categorical cross-entropy, Jaccard loss (1 − intersection over union) and Kullback–Leibler divergence and their weighted combinations were tested. Because the segmentation results were similar for the tested loss functions, only the results for the categorical cross-entropy loss function are presented in this paper. Customized loss functions including terms to maximize the peakiness of a single prediction, and on the other hand to maximize the spread of the classes over each training batch were also studied shortly. These additional loss terms were combined with the above-mentioned semantic segmentation loss functions. To achieve useful results with spread maximization included, a rather large batch size is required. This results to long training times with the current hardware setup. The approach using a combination of loss functions as its loss function will still require adjustment, i.e. finding the most suitable fractions for the loss terms. The Adam optimizer (Kingma and Ba, 2014) was applied in all the training variations. The initial learning rate applied was 0.0001. In total, 150 epochs were run and the weights were selected to represent the epoch with the minimum validation loss. Typically this minimum was achieved after 20–40 epochs.

The hardware used in the study was a common desktop computer with Ubuntu 20.4 Linux operating system. The Central Processing Unit (CPU) was an AMD Ryzen 5 2400G with eight cores and with AMD Radeon Vega Graphics and 16 GB of
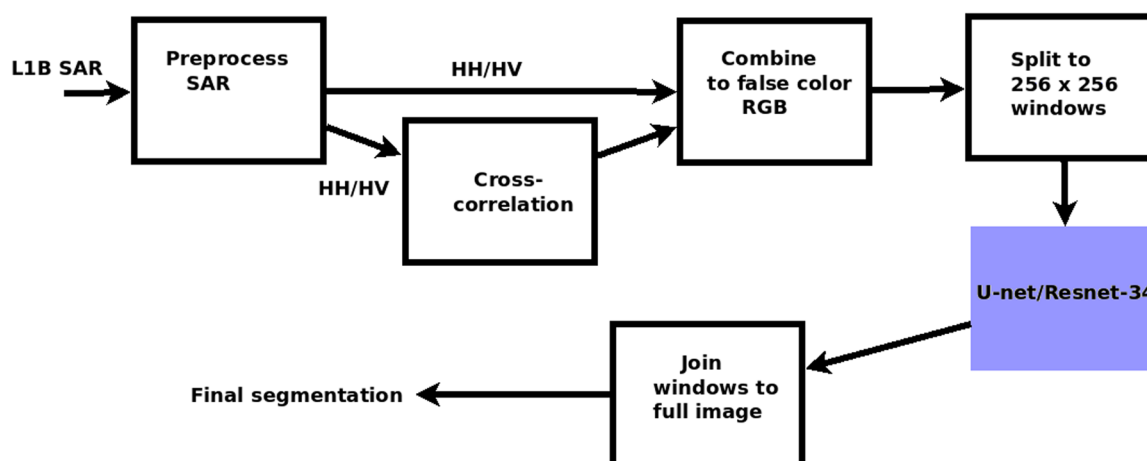

Figure 7. Diagram of the segmentation phase. L1B SAR data preprocessing is similar as in training phase.

RAM. The training time for the whole training dataset was a few hours using CPU. The AMD Graphics Processing Unit (GPU) was not supported by the python library used. Segmentation execution times for a daily mosaic in the 500 m resolution are 1–2 min on the same CPU, thus making this segmentation approach suitable for operational purposes. In the future also higher resolutions will be considered in operational use, e.g. for a 100 m resolution mosaic the execution time required for segmentation would increase to ∼0.5–1 h in this particular hardware setup (execution time increase can be estimated to be linear because the images are processed as cropped smaller windows sequentially) but utilizing more efficient hardware, e.g. with multiple better suitable (NVIDIA) GPUs or updated libraries for AMD GPUs, would decrease the training and segmentation execution times significantly, and the method will then be suitable for operational use also in a higher resolution than the resolution used in this study.

4. Results

As already mentioned in Section 2, in practice the segmentation produced only ten classes containing a reasonable amount of data. The other classes, based on quantization of Eqn (1), were very rare or non-existent in the actual training and test datasets. Only a few pixels of some of these theoretical classes in the test dataset were detected, so only the actually occurring ten classes have been used and reported in the following results.

If we make a direct inter-class comparison between the classes in the FMI ice charts and the U-net/ResNet-34 segmentation, the correspondence is not good, only open water class has a good correspondence. However, as the major objective is to perform segmentation, not classification, a confusion matrix, for example, is not a suitable measure to measure the performance for this non-semantic segmentation.

If the classes are reduced to two classes of open water and sea ice, then 94% of the water is correctly classified and 73% of the sea ice is correctly classified by the U-net/ResNet-34 algorithm. For the melt period the statistics was even better: 98% of the open water were correctly classified and 90% of the sea ice were correctly classified. During the freeze-up period (late December—mid-January) some thin level ice areas were incorrectly classified to open water by the algorithm, decreasing the sea-ice classification performance to 70% during freeze-up. Also later during the ice season some new ice zones (according to the ice charts) were misclassified to open water. However, these numbers just describe the performance of classification into these two classes but they are not a good measure of the segmentation as a whole. For example, if there appear small open water segments within an area that actually is sea ice, these segments can be classified to sea ice in the separate classification phase after the segmentation, thus correcting the possible preliminary misclassification made by the segmentation algorithm.

Segment sizes of the different segmentation methods were compared. The results are shown in Table 2.

The numbers clearly indicate that the segment sizes for U-net/ResNet-34 are significantly larger than for the operational ICM

segmentation. Especially for open water, the average segment sizes are larger, indicating that the ICM segment fragmentation over open water is significantly reduced. The average segment sizes over sea ice are also larger for U-net/ResNet-34 than for ICM, but they still are smaller than ice chart polygons.

To compare the correspondence of segments between the FMI ice charts (polygons) and both the operational ICM segmentation and the U-net/ResNet-34 segmentation, the intersection over union (IoU) metric was used. The IoU to compare two segments or classes is defined as:

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}, \tag{2}$$

where A and B are the segmentation results, or in this case of FMI ice charts polygons and segmentation results, to be compared. By $|X|$ the area of X is denoted, here the areas correspond to the pixel counts of the segment or polygon area. In this study IoU was computed between the ice chart polygon classes and classes produced by the segmentation and also between individual ice chart polygons and segments. IoU between ice chart polygons and segments is denoted by IoU_s and IoU between ice chart and polygon classes by IoU_c in the following. In IoU_c the numerator part also counts the matching pixels (same class) of a certain class in two images, and the denominator part counts the pixels that belong to the certain class in either of the two images. This computation is performed for all the existing classes. In IoU_s the numerator part is the number of pixels belonging to both the ice chart polygon and segment to be compared and the denominator part is the number of pixels belonging to either the ice chart polygon or the segment. IoU is a number between zero and one, and the value of one indicates perfect correspondence between the two segments in comparison. Here, IoU is given in percents, i.e. the IoU range here is 0–100. The IoU_c values were computed after mapping the ICM and U-net/ResNet-34 segments to the ice chart segment classes by applying a majority voting for each segment. This approach enables comparison of the segmentation results with respect to the rasterized ice chart polygons. The results of this comparison are shown in Table 3.

The IoU_c for open water was 94% for the U-net and 53% for sea-ice classes (average weighted by frequency of occurrence). The corresponding values for ICM were 94 and 54%, respectively. The accuracy based on IoU_c is actually very similar for ICM and U-net/ResNet-34. The results indicate that the correspondence of the segments using ICM is slightly better than for U-net/ResNet-34. This difference can be explained by the significantly smaller segment size provided by ICM. These many small segments are able to cover the ice chart polygons in more detail than the larger segments produced by U-net/ResNet-34. However, in many

Table 2. Average segment size for open water (OW) and sea-ice (SI) areas for ice charts (average polygon size), ICM segmentation and U-net/ResNet-34 segmentation

Source	OW	SI
Ice chart	39 083	698
ICM	274	132
U-net	4692	397

The areas are given in km².

Table 3. U-net/ResNet-34 and ICM segmentation class-wise IoU_c scores for the used sea-ice classes compared to the ice chart polygons

Class	U-net/ResNet-34 IoU$_c$	ICM IoU$_c$	Proportion (all)	Proportion (SI)
1	93.7	94.2	73.9	–
8	47.6	52.2	3.6	13.8
9	40.0	37.9	2.1	8.0
10	25.7	30.1	1.1	4.2
11	54.1	51.5	2.5	9.6
18	41.3	48.5	1.0	3.8
24	59.2	64.1	3.3	12.6
26	75.7	85.0	3.1	11.9
28	63.7	77.7	5.0	19.2
32	39.6	11.2	4.4	16.9
All	87.8	89.7	100.0	100.0

Also the proportions of the classes, based on the ice chart segments, are given both for the whole area (all) and sea-ice (SI) areas separately.

Table 4. U-net/ResNet-34 and ICM segmentation IoU_s scores for the best matching single segments, winter refers to the whole winter with the melting period excluded and melt refers to the melting period

Class	U-net/ResNet-34 IoU_s	ICM IoU_s
winter SI	30.3	16.5
winter OW	92.6	9.6
melt SI	29.3	18.5
melt OW	90.7	3.1

SI, sea ice; OW, open water.

cases, assigning ice parameters to small segments is in practice difficult for both human and machine because of limited available contextual information within the small segments. U-net/ResNet-34 provides a more general view with larger segments but still a good correspondence between ice chart segments. The total correspondence was almost 90% for both the segmentation methods. This is because the match for classes covering most of the area (1, 8, 24, 26, 28 and 32) was higher. For the classes covering less area (pixels) the matches were not that good.

The other measure, IoU_s, to compare the segmentation methods and the ice chart polygons was to locate the best matching segment, i.e. the segment with the largest IoU, corresponding to each ice chart polygon and to compute the average IoU_s of the best matched for all the ice chart polygons. This measure is completely independent of the class of the segments, only segments are compared with respect to each other. The measure was computed for the sea-ice segments and open water areas separately. The values were also computed for the melting period and for the winter with the melting period excluded separately. These results are shown in Table 4. The results were rather similar to the other way around, i.e. IoU_s for ice chart polygons best matching each segment produced by a segmentation algorithm.

Based on these statistics it is evident that the correspondence between the ice chart polygons and the segments is better for the U-net/ResNet-34 than for ICM, for both sea ice and for open water. Especially, for open water areas U-Net/ResNet-34 seems to perform very well. These measures for U-net/ResNet-34 can possibly be improved by using a larger training dataset and by balancing the class proportions in the training. However, complete correspondence between ice chart polygons and SAR segments using these classes provided by the U-net/ResNet-34 is not actually possible because, in ice charting, complementary information from other EO sources and in situ

observations are utilized. Ice analysts also have the knowledge of the ice development in the past. To include the past ice development information in the form of a SAR mosaic time series, instead of single daily SAR mosaics, in the segmentation is also possible in future algorithm development.

In Figure 8, the current operational FMI segmentation results, using a combination of MS clustering and ICM contextual segmentation, for the 15th day of January–April 2021 SAR mosaics are shown. In the figure, the number of clusters varies from one image to another, the scale just identifies the increase in the total σ^0 magnitude $|\sigma^0| = \sqrt{(\sigma^0_{HH})^2 + (\sigma^0_{HV})^2}$. It can be seen that ICM segmentation produces quite many separate segments in areas that actually represent the same kind of target. Especially, this can be seen in the open water areas where many irrelevant segments are produced. In the areas of σ^0 varying as a function of the range (incidence angle) this can be seen as a ramp-like behavior with many jagged segment boundaries in the direction perpendicular to the range direction. This kind of oversegmentation sometimes also happens over the smooth level ice. However, this phenomenon is less prominent over sea ice than over open water.

Compared to ICM segmentation, the U-net/ResNet-34 segmentation in Figures 9E–H behaves locally in a different way: uniform areas, especially open water appears as single segments and the segment boundaries correspond well to visual interpretation of the imagery. This was the major objective of this study and the U-net/ResNet-34 segmentation model seems to provide useful results from this point of view.

It should be noted that the segment color coding in Figures 8 and 9 cannot be directly compared. Instead, one should pay attention to the segment boundaries and differences in them. This kind of a visual comparison reveals many clear differences: the ICM segmentation produces many more segments and the segment boundaries are not always natural because of varying σ^0 over the image due to different scattering conditions (different local SAR viewing angles) from similar target surfaces. In the studied U-net/ResNet-34 segmentation, this effect is typically not present, making further visual and automated interpretation of the local ice conditions significantly easier.

In Figures 10–12 the U-net/ResNet-34 segments are presented with classes mapped to those of the same day ice chart grid. This makes visual comparison of the segments and their boundaries easy. Some details of the segmentation results are
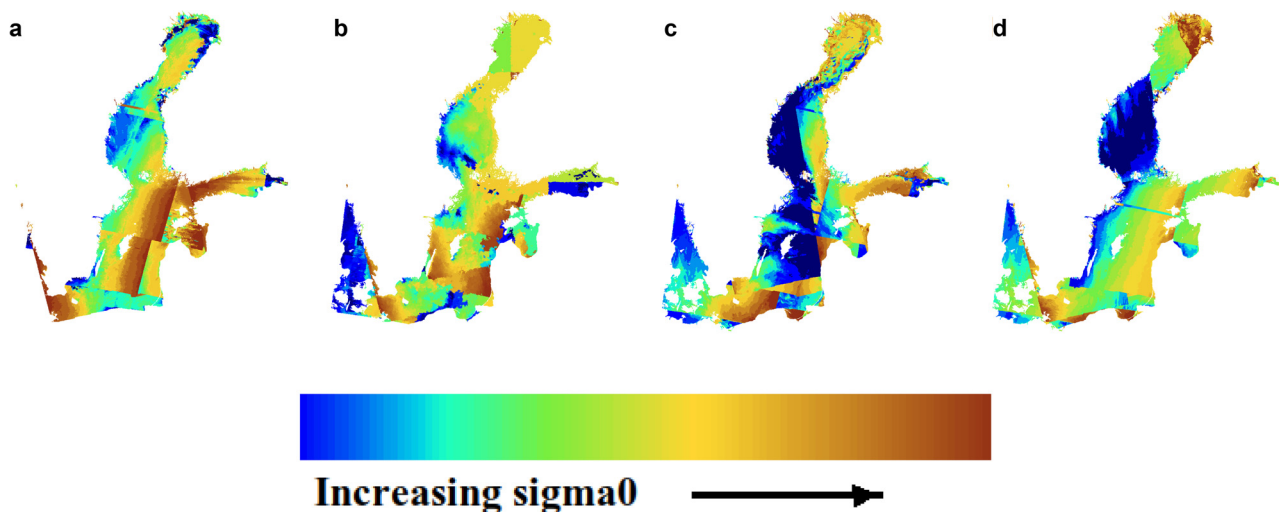

Increasing sigma0

Figure 8. Segmentation results of the 15 January (A), February (B), March (C) and April (D) 2021 image mosaics by the FMI operational method, based on MS and ICM. Different mosaics have different number of classes produced by the MS clustering.
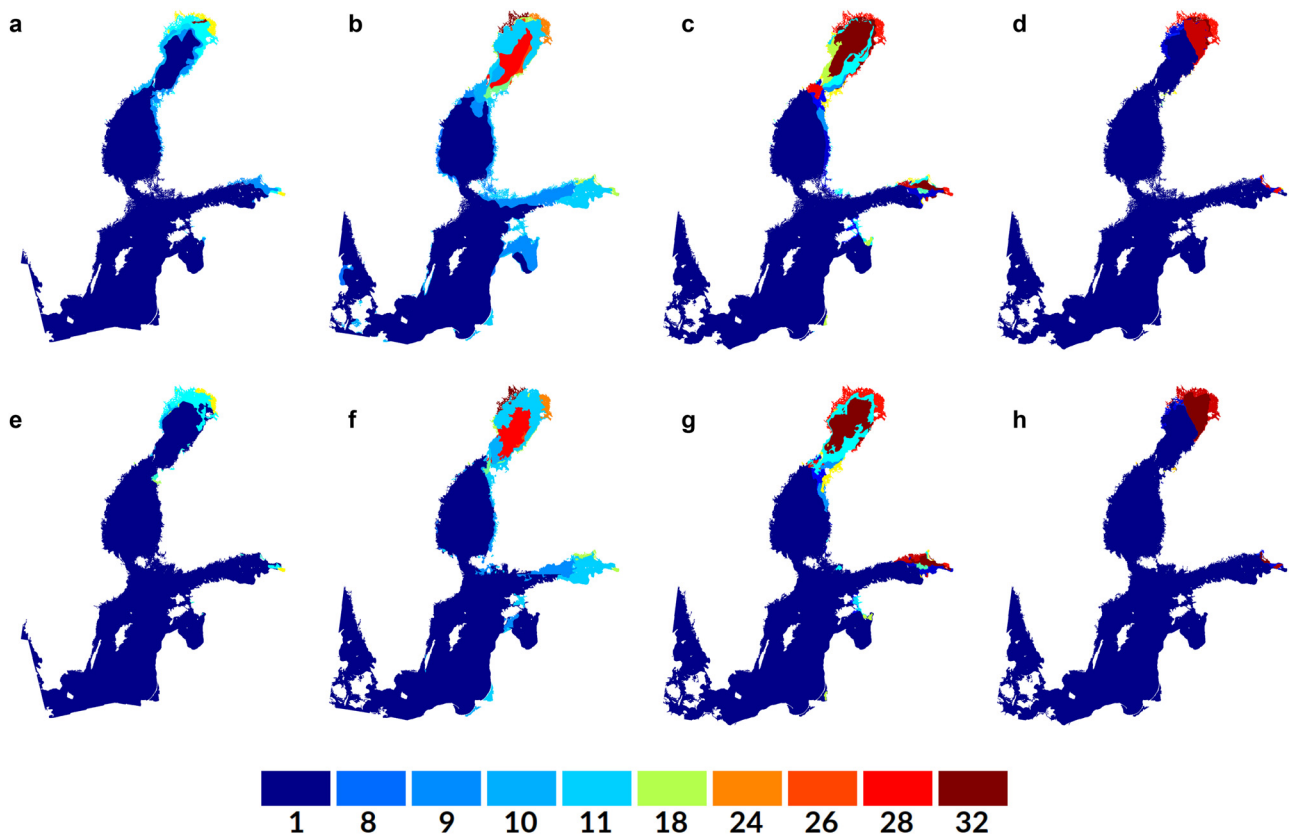
Figure 9. Classes of the 15 January, February, March and April 2021 FMI ice charts (a–d) and the corresponding U-net/ResNet34 segmentation results (e–h).

shown in Figures 10–12. The ice chart polygons are also shown in the figures. For visual reference, also the cropped daily SAR image mosaic images for HH and HV channels have been included in the figures. In Figure 10 an eastern Gulf of

Finland segmentation corresponding to a mid-winter SAR mosaic is shown. It can be seen that the ice chart polygons and segments correspond to each other rather well. In Figure 11 a Gulf of Bothnia, the northernmost part of the
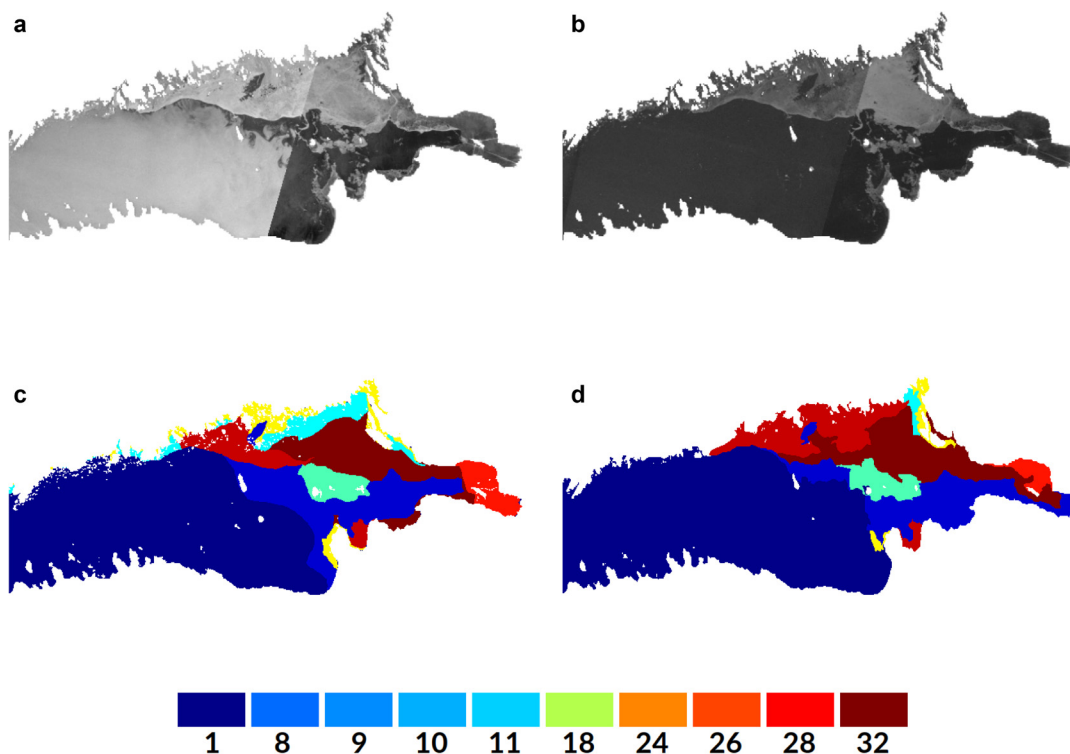

Figure 10. A detail over eastern Gulf of Finland, 15 March 2021, SAR mosaic HH channel (a), SAR mosaic HV channel, ice chart polygons (c) and U-net/ResNet-34 segments mapped to ice chart classes (d).
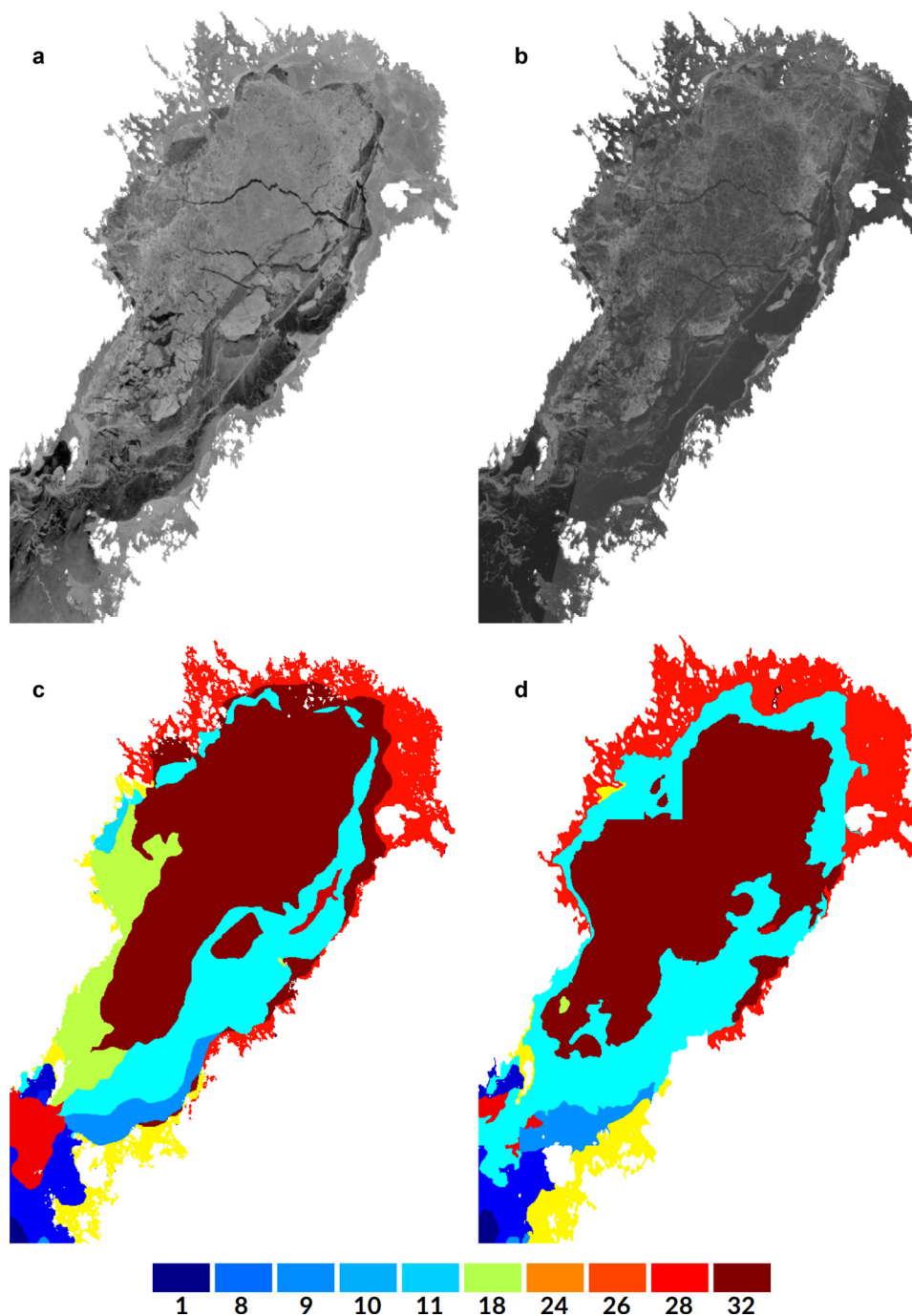
Figure 11. A detail over Gulf of Bothnia, 15 March 2021, SAR mosaic HH channel (a), SAR mosaic HV channel (b), ice chart polygons (c) and U-net/ResNet-34 segments mapped to ice chart classes (d). A processing artifact can be seen in the northwestern part of the large class 32, the class boundary seems to follow the processing block boundary.

Baltic Sea, segmentation corresponding to the same mid-winter SAR mosaic is shown. In this case there exist some more differences between the ice chart polygons and segments. However, still for example the land fast ice (representing class 28 in the figure) is well distinguished in the segmentation result. It is noticeable that a computation block boundary is visible in the northwestern part of the large class 32 area in this figure. This kind of artifacts were rare in the classification results but still exist in some cases in the daily test data classification results, covering the whole winter 2020–2021. The third detail example of Figure 12 is over the Gulf of Bothnia during the melting period and with wet snow on the sea ice. Due to the wet surface the SAR backscattering from the sea ice was reduced and contrast between sea ice and open water was also reduced.

Still, the sea ice and open water segments were well distinguished from each other by the U-net/ResNet-34 segmentation.

5. Discussion and conclusions

In this study, U-net/ResNet-34 segmentation was applied to Baltic Sea SAR imagery. According to the experiments performed, it was evident that the U-net/ResNet-34 segmentation is a suitable candidate for automated sea-ice segmentation and capable of producing sea-ice information over wide areas of the Baltic Sea in a compact form. In the following also some possible future research directions will be discussed. The focus of the proposed future research will be in further development of the proposed method and fine-tuning it for operational use. One important objective
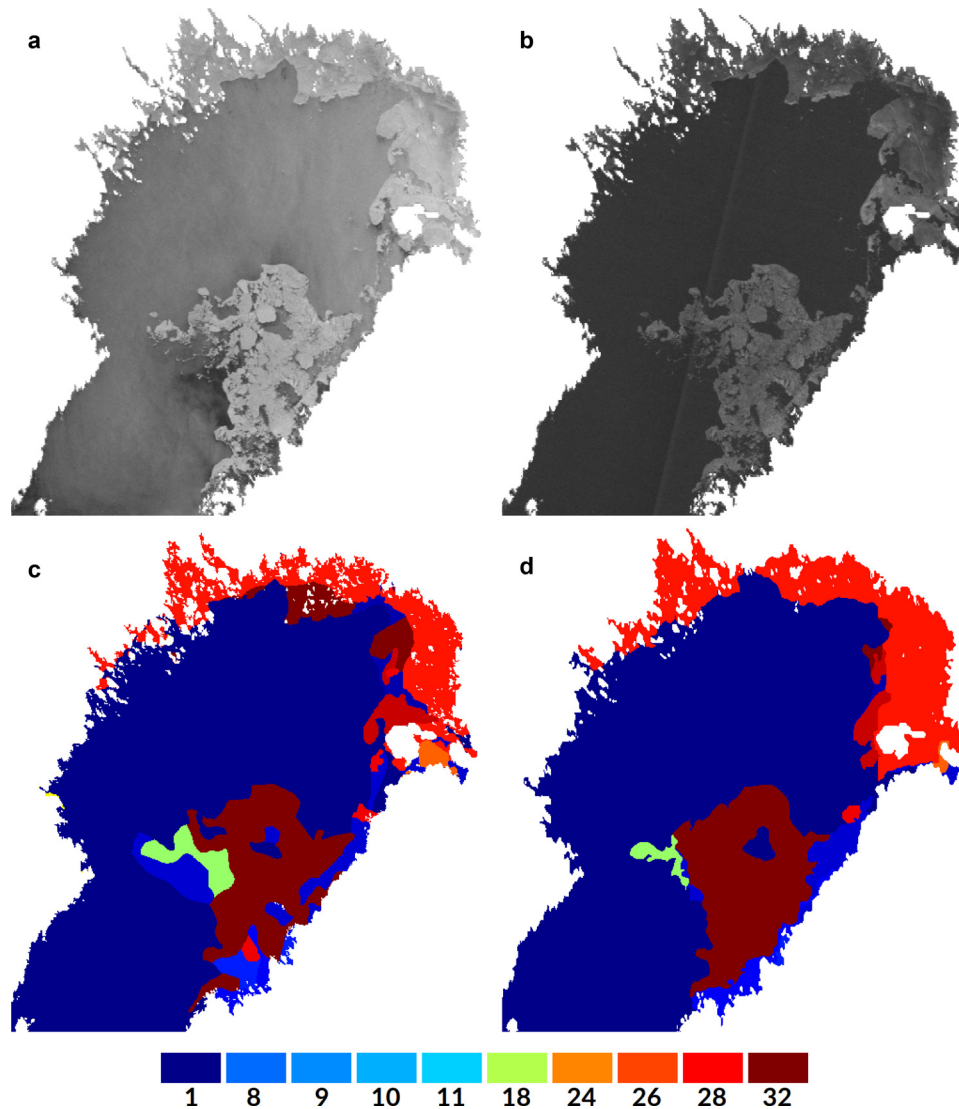
Figure 12. A late melting period detail over Gulf of Bothnia, 28 April 2021, SAR mosaic HH channel (a), SAR mosaic HV channel (b), ice chart polygons (c) and U-net/ResNet-34 segments mapped to ice chart classes (d).

will be integration of the new segmentation method in to the FMI operational sea-ice production chain providing segment-wise sea-ice products for FMI internal use and the Copernicus Marine Service. Another objective will be to provide supporting information for ice charting at the Baltic Sea ice service.

Visual inspection indicated reasonable results regarding the segmentation only, not the segment classes. This was supported by the IoU metrics (IoU_c and IoU_s) providing information on the correspondence of the segments and ice chart polygons. Because the objective of this study was to use the result for segmentation only, the classification, provided by semantic segmentation in addition to the segmentation, is not essential in the context of this study. The ability to provide reasonable segments to be used in segment-wise ice characterization is the most important result.

One objective of the study was to reduce the oversegmentation provided by the current FMI operational ICM segmentation, especially over open water areas. For successful segment-wise ice parameter estimation, a segment should contain a sufficient amount of pixels for estimating the sea-ice properties (distribution if the segment is not homogeneous) within the segment. Based on the computed segment size statistics, the U-net/ResNet-34 approach seems to produce significantly larger segments than the ICM segmentation and also the segment

boundaries rather well correspond to the boundaries of the different ice fields and the ice chart polygons.

The performed study also indicated that approximately ten classes are a suitable number of classes for Baltic Sea ice SAR segmentation. This estimate is based on the performed tests: the number of classes originally trained (based on the quantization of the three independent ice chart parameters) was larger but after the training the neural network provided only this reduced set of classes and the remaining classes did not appear in the semantic segmentation outputs. Also, the segmentation results using this number of classes were reasonable and the major ice and open water areas were separated. The ice classes used were generated based on quantization of three polygon-wise quantities of the digitized FMI ice charts. The basic idea was to produce sea-ice classes describing the navigability in ice. Also other selection of classes can be considered in the future research. One possibility is to cluster the ice chart polygons based on the polygon-wise SAR properties (σ^0 and texture) and use the clustering result as classes for the training, possibly in a semi-supervised manner. Also, a two-stage classification could be used: first distinguish sea ice and open water segments, and then perform a separate segmentation for sea-ice areas by another neural network. The advantage of this two-stage approach would be to exclude the open water areas with a large variety of σ^0 from the second

stage to make the clustering more robust. U-net/ResNet-34 segmentation, with the training dataset used in this study, seems to take into account the SAR σ^0 variation caused by different incident angles. The effect of incidence angle is visible more prominently over open water areas and the segmentation seems to distinguish these areas well and also map them to the same semantic class. An obvious reason for this are the texture patterns of the open water class in the training dataset learned by the algorithm, not only σ^0 magnitude.

It is also quite straightforward to include local incidence angle as one additional channel in the segmentation but in the light of the obtained results this does not seem to be necessary. The open water areas with the strongest incidence angle dependency were well detected by the algorithm without the incidence angle information and over sea ice incidence angle dependence was not significant after the linear incidence angle correction, based on the incidence dependencies defined in Makynen and Karvonen (2017), performed.

The training dataset used in this study was randomly sampled and was not balanced between the classes, i.e. in the training dataset the proportions of the classes corresponded to their proportions in the Baltic Sea during the season 2018–2019. The open water class had the largest proportion of the classes for the season. One topic for future research is to test training with a balanced dataset. Also use of other class proportions different from their true seasonal occurrence could be tested to find an optimal training dataset. The U-net/ResNet-34 segmentation trained with the current training dataset, even with its limitations, indicated that it has potential in segmentation of the operational FMI SAR imagery over the Baltic Sea.

The segmentation provided by the U-net/ResNet-34 segmentation is clearly able to distinguish between different ice and open water or very low ice concentration areas. For example, land-fast ice areas, level ice areas and drift ice areas can be distinguished to separate segments by the method. At this point (segmentation without tightly fixed segment classes) of the automated sea-ice information processing chain, the segment labels are not essential. Only the ability of the algorithm to generate segments corresponding to uniform sea-ice areas is essential. The more detailed ice information will be assigned to the segments in a later analysis phase. The analysis phase can be automated or manually performed by ice analysts. Examples of applying a segmentation first and assigning SIC to the segments were proposed, e.g. in Karvonen (2017, 2021). The next processing phase after the segmentation will be to perform a thorough automated sea-ice analysis over each segment using EO data from multiple sources, including SAR, microwave radiometer (AMSR2) and radar altimeter (CryoSat-2 SIRAL or Sentinel-3 SRAL) data. One useful piece of information for navigation is the Risk Index Outcome (RIO) (IMO, 2016). RIO is dependent on the ship ice class and thus different for different vessels. RIO provides a ship the information about which areas it is safe to sail with the particular ship. Segmentwise RIO estimates based on ice information derived from remote-sensing data is one potential parameter to be provided to aid navigation.

It is also noteworthy that the image mosaics used for testing the segmentation algorithm consisted of both calibrated Sentinel-1 and Radarsat-2 images acquired approximately at the same frequency and polarization combination. The segmentation seemed to work well for these mosaics even though only Sentinel-1 imagery were used in the training phase. Compatibility between Sentinel-1 and Radarsat-2 data has also been observed in Guo and others (2022), with some exceptions due to the different noise floors of the two instruments. Radarsat Constellation Mission (RCM) data are acquired at the same frequency and training data compatibility between Sentinel-1 and RCM data will be one interesting topic for future studies.

This kind of segmentation can also be used as an initial state of manual ice charting. The segments can be converted into polygons in a vector graphics format, such as ESRI shape files (ESRI, 1998), used by many ice services. The ice analysts can then edit the segments converted to polygons in their ice charting software and assign detailed ice information to them.

The input segmentation for daily ice charting should not include too many details because the ice analysts must have the ice charts ready by a fixed time every afternoon. Too detailed inputs would require too much manual work to get completed within the strict time limit.

In the late melting period even distinguishing between open water and level ice with liquid water or very wet snow on it may become difficult by SAR and this kind of ice condition may provide misinterpretation by both automated algorithms and visual inspection if the segmentation or polygonalization is based on SAR data alone. One topic of future research is to study whether including data from microwave radiometer (e.g. as additional image channels) would improve the segmentation in melting period ice conditions with a wet or refrozen ice surface. For the test dataset of the 2020–2021 season mosaics, the U-net/ResNet-34 algorithm seemed to perform well during the melt period. The melt period started already in early April 2021. Some areas classified as thin level ice in ice charts was mapped to open water by the segmentation algorithm. The SAR backscattering from these areas is very similar to backscattering from open water, i.e. speckle noise without any features, and distinguishing is very difficult. The ice analysts typically have some additional information (optical/infrared image data, microwave radiometer data, X-band SAR data and knowledge of the ice development history) at their disposal, compared to the segmentation algorithm. The ice development history could be taken into account in an automated algorithm by including a time series of SAR mosaics of the days before the segmentation mosaic date in the segmentation algorithm as additional image channels. This will be one topic for future research. To be able to confirm and optimize the algorithm performance of distinguishing between level ice and open water in freeze-up and melt conditions, a large multi-year representative dataset covering different freeze-up and melt period situations will be required. This will also be one topic of future research.

The future work will also include research of applying machine learning methods to extract sea-ice information from multiple data sources available in near-real-time, including ice modeling to complement information obtained from EO data, and assigning these information to each segment produced by the segmentation algorithm. One interesting topic for future research would also be to apply explainable AI (XAI) (Arrieta and others, 2013) to analyze dependencies between the inputs, neural network internals (feature learning/extraction) and outputs to get a better insight of the neural network functioning.

More detailed information on sea ice within segments can be included in a final sea-ice product as separate layers. For example, leads, cracks and large pressure ridges distinguishable in SAR imagery or their segment-wise statistics can be included. For this purpose, specific approaches to get detailed fine-scale information should be used; one alternative is to utilize weakly supervised deep learning (Wang and others, 2020) to extract details from SAR imagery.

One possible way to improve the segmentation performance would be to utilize an adversarial network (Goodfellow and others, 2014) in optimizing the segmentation. If more detailed segments would be required, then one way to enable more details would be requiring in the training phase that, e.g. polygon/segment class modes, instead of pixel-wise polygon class values, agree with the ice chart polygon class. Another factor effecting the segmentation and classification is the selection of the loss function. In this study, some common semantic segmentation loss functions were tested. They had very similar segmentation performance for the test dataset.

Also tests with a customized loss function were performed. According to preliminary tests, it seems possible to adjust the level of segmentation detail, at least to some extent, by using such customized loss function. A thorough comparison of loss functions, including loss functions that are combinations of more than one loss function, and estimation of optimal weights for different combinations of loss functions for the segmentation task will be a laborous but important part of the future development. Also, use of novel loss functions especially tuned for the segmentation task should be studied more.

Also, using a constant image grid without data augmentation is an interesting topic of future research. It can be assumed that the neural network will learn the typical local ice patterns, especially near coastline because coastline is similar in each image block of the same location. This approach will require SAR mosaics in the fixed grid over multiple ice seasons to be able to train the typical local ice characteristics.

At FMI there are also plans to test the application of the proposed U-net/ResNet-34 segmentation algorithm to Arctic Ocean SAR data, first using the SAR mosaics produced over a test area in Barents and Kara Seas. These data have been archived at FMI for a time period of several years. The Arctic segmentation may require additional training data. For example, the Russian Arctic-Antarctic Research Institute (AARI) ice charts (http://wdc.aari.ru/datasets/) could be used for training over the Arctic test area. Unfortunately, the Arctic ice charts are less accurate then the Baltic Sea ice charts. They only contain information of SIC and stage of ice development and the polygons in them are rather large, providing only coarse scale sea-ice information, i.e. they contain significantly less details compared to the gridded Baltic Sea ice charts. One approach at least partially to overcome this deficiency and to be tested will be to use Baltic Sea data for training and after that applying the algorithm trained using the Baltic Sea data to the Arctic SAR data. This approach may require some additional adjustment because in the Arctic there occur ice types not existing in the Baltic Sea. For example, detection of multi-year ice segments should be based on Arctic training data. One possibility is to apply a two-stage segmentation, one based on an Arctic training and the other based on a Baltic Sea training if the Baltic Sea training is seen to provide any added value.

In conclusion, the studied algorithm is a good candidate for operational SAR segmentation at FMI, also as part of the Copernicus Marine Service production chain. It produced useful results for an independent test dataset, even with a rather limited training dataset and the execution time is not too long for operational use. However, before replacing the operational FMI SAR segmentation algorithm by U-net/ResNet-34 training with a large representative training dataset, covering multiple winter seasons and fine tuning of the algorithm parameters based on the large and representative dataset will be needed. This dataset can be composed of the Baltic Sea Sentinel-1 and Radarsat-2 images and daily ice charts archived at FMI since the winter season 2014–2015.

References

Al Bayati M and El Zaart A (2013) Automatic thresholding techniques for SAR images, In Proceedings of International Conference in Soft Computing, pp. 77–84, doi: 10.5121/csit.2013.3308

Aleissaee AA and 6 others (2021) Transformers in remote sensing: a survey. *Remote Sensing* 15, 1860. doi: 10.3390/rs15071860

Arrieta AB and 11 others (2013) Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *An International Journal on Information Fusion* 58(3), 82–115. doi: 10.1016/j.inffus.2019.12.012

Ayed IB, Vazquez C, Mitiche A and Belhadj Z (2004) SAR image segmentation with active contours and level sets, In Proceedings of 2004 International Conference on Image Processing, 2004. ICIP '04, Vol. 4, Singapore, 2004, pp. 2717–2720, doi: 10.1109/ICIP.2004.1421665

Besag JE (1986) On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society, Series B* 48(3), 259–302.

Boulze H, Korosov A and Brajard J (2020) Classification of sea ice types in Sentinel-1 SAR data using convolutional. *Neural Networks Remote Sensing* 12(13), 2165. doi: doi.org/10.3390/rs12132165

Bourbigot M, Johnsen H and Piantanida R (2016) Sentinel-1 product specification, ESA, S1-RS-MDA-52-7441, 2016.

Brown PO and 7 others (2022) Mahalanobis distance based K-means clustering. In Big Data Analytics and Knowledge Discovery: 24th International Conference, DaWaK 2022, Vienna, Austria, August 22–24, 2022, Proceedings. Springer-Verlag, Berlin, Heidelberg, pp. 256–262. doi: 10.1007/978-3-031-12670-3_23

Cheng Y (1995) Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(8), 790–799. doi: 10.1109/34.400568

Deng H and Clausi DA (2005) Unsupervised segmentation of synthetic aperture radar sea ice imagery using a novel Markov random field model. *IEEE Transactions on Geoscience and Remote Sensing* 43, 528–538. doi: 10.1109/TGRS.2004.839589

Dosovitskiy A and 6 others (2021) Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N (2020) An image is worth 16×16 words: transformers for image recognition at scale, ICLR 2021, cite arxiv:2010.11929.

Dowden B, De Silva O, Huang W and Oldford D (2020) Sea ice classification via deep neural network semantic segmentation. *IEEE Sensors Journal* 21(10), 11879–11888. doi: 10.1109/JSEN.2020.3031475

Environmental Systems Research Institute Inc. (1998), ESRI shapefile technical description, 1998, online: https://www.esri.com/content/dam/esrisites/sitecore-archive/Files/Pdfs/library/whitepapers/pdfs/shapefile.pdf.

Ester M, Kriegel HP, Sander J and Xu X (2006) A density-based algorithm for discovering clusters in large spatial databases with noise, United States: N. p., In *Proceedings of the second international conference on knowledge discovery & data mining*, Simoudis, E, Han, J, Fayyad, U (eds), pp. 226–231, online: https://www.osti.gov/biblio/421283.

Goodfellow I and 7 others (2014) Generative adversarial nets. In Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014), pp. 2672–2680.

Granskog M, Kaartokallio H, Kuosa H, Thomas DN and Vainio J (2006) Sea ice in the Baltic Sea – a review. *Estuarine Coastal Shelf Science* 70, 145–160.

Guo W, Itkin P, Lohse J, Johansson M and Doulgeris AP (2022) Cross-platform classification of level and deformed sea ice considering per-class incident angle dependency of backscatter intensity. *The Cryosphere* 16, 237–257. doi: 10.5194/tc-16-237-2022

He K, Zhang X, Ren S and Sun J (2016) Deep residual learning for image recognition, In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90

Hochreiter S (1991) Untersuchungen zu dynamischen neuronalen Netzen, Diplom thesis. Institut f. Informatik, Technische Univ. Munich, In German.

Huang Y, Ren Y and Li X (2021) Classifying sea ice types from SAR images using a U-net-based deep learning model. In IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, pp. 3502–3505. doi: 10.1109/IGARSS47720.2021.9554511

Iakubovskii P (2019) Segmentation models, GitHub repository, publisher GitHub, 2019, online: https://github.com/qubvel/segmentation_models.

International Maritime Organization (IMO) (2016), Guidance on methodologies for assessing operational capabilities and limitations in ice. MSC.1/Circ 1519, June 6, 2016. Online: https://www.nautinst.org/uploads/assets/uploaded/2f01665c-04f7-4488-802552e5b5db62d9.pdf.

JCOMM (Joint World Meteorological Organization (WMO) Intergovernmental Oceanographic Commission (IOC) Technical Commission for Oceanography and Marine Meteorology) (2014) SIGRID-3: a vector archive format for sea ice charts, JCOMM-TR-023, revision 3, available at: https://www.jcomm.info/index.

php?option=com_oe&task=&viewDocumentRecord&docID=4439 (last accessed: 1 February 2023).

Jeevitha K, Iyswariya A, RamKumar V, Basha SM and Kumar V-P (2020) A review on various segmentation techniques in image processing. *European Journal of Molecular & Clinical Medicine* **7**(4), 1342–1348.

Kankaanpaa P (1997) Distribution, morphology and structure of sea ice pressure ridges in the Baltic Sea. *Fennia* **175**, 139–240.

Karvonen J (2004) Baltic Sea ice SAR segmentation and classification using modified pulse-coupled neural networks. *IEEE Transactions on Geoscience and Remote Sensing* **42**(7), 1566–1574. doi: 10.1109/TGRS.2004.828179

Karvonen J (2015) Evaluation of the operational SAR based Baltic Sea ice concentration products. *Advances in Space Research* **56**, 119–132. doi: 10.1016/j.asr.2015.03.039

Karvonen J (2017) Baltic Sea ice concentration estimation using Sentinel-1 SAR and AMSR2 microwave radiometer data. *IEEE Transactions on Geoscien* **55**(5), 2871–2883. doi: 10.1109/TGRS.2017.2655567

Karvonen J (2021) Baltic Sea ice concentration estimation from C-band dual-polarized SAR imagery by image segmentation and convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing* **60**, doi: 10.1109/TGRS.2021.3097885

Karvonen J, Simila M and Heiler I (2003) Ice thickness estimation using SAR data and ice thickness history, IGARSS 2003. In 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477), Toulouse, France, 2003, vol. 1, pp. 74–76. doi: 10.1109/IGARSS.2003.1293683

Karvonen J, Vainio J, Marnela M, Eriksson P and Niskanen T (2015) 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* **8**(4), 1799–1807. doi: 10.1109/JSTARS.2015.2426414

Kingma DP and Ba J (2014) Adam: a method for stochastic optimization. URL: http://arxiv.org/abs/1412.6980.

Lee J-S and Jurkevich I (1988) Segmentation of SAR images. *IEEE Transactions on Geoscience and Remote Sensing* **27**(6), 674–680. doi: 10.1109/36.35954

Li W and 6 others (1989) Watershed-based hierarchical SAR image segmentation. *International Journal of Remote Sensing* **28**(17), 3377–3390. doi: 10.1080/014311699211390

Li Y, Li H, Fan D, Li Z and Ji S (2023) Improved sea ice image segmentation using U2-net and dataset augmentation. *Applied Sciences* **13**, 9402. doi: 10.3390/app13169402

MacQueen JB (1967) Some methods for classification and analysis of multivariate observations. In Le Cam LM and Neyman J (eds), *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Vol. 1, California: University of California Press, pp. 281–297.

Maillard P, Clausi DA and Deng H (2006) Operational map-guided classification of SAR sea ice imagery. *IEEE Transactions on Geoscience and Remote Sensing* **43**(12), 2940–2951. doi: 10.1109/TGRS.2005.857897

Makynen M and Karvonen J (2017) Incidence angle dependence of first-year sea ice backscattering coefficient in Sentinel-1 SAR imagery over the Kara Sea. *IEEE Transactions on Geoscience and Remote Sensing* **55**, 6170–6181. doi: 10.1109/TGRS.2017.2721981

Malmgren-Hansen D and Nobel-Jorgensen M (2015) Convolutional neural networks for SAR image segmentation, In Proc. 2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT). doi: 10.1109/ISSPIT.2015.7394333

MANICE (2005), Manual of standard procedures for observing and reporting ice conditions, Revised Ninth Edition, Canadian Ice Service, Ottawa, Ontario, Canada, June 2005.

Manos E, Witharana C, Udawalpola MR, Hasan A and Liljedahl AK (2022) Convolutional neural networks for automated built infrastructure detection in the Arctic using sub-meter spatial resolution satellite imagery. *Remote Sensing* **14**(11), 2719. doi: 10.3390/rs14112719

MDA Ltd (2018) Radarsat-2 product description, Report RN-SP-52-1238, Maxar Technologies Ltd., available online: https://earth.esa.int/eogateway/documents/20142/0/Radarsat-2-Product-description.pdf.

Oliver CJ (1994) Edge detection in SAR segmentation, In Proc. SPIE 2316, SAR Data Processing for Remote Sensing, doi: 10.1117/12.197528

Ren J, Hwang B, Murray P, Sakhalkar S and McCormack S (2015) Effective SAR sea ice image segmentation and touch flow separation using a combined multi-stage approach. In *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Milan, Italy, pp. 1040–1043, doi: 10.1109/IGARSS.2015.7325947

Ren Y, Li X, Yang X and Xu H (2023) Sea ice detection from SAR images based on deep fully convolutional networks. In Li X and Wang F (eds), *Artificial Intelligence Oceanography*. Singapore: Springer. doi: 10.1007/978-981-19-6375-9_12

Ristea N-C, Anghel A and Datcu M (2023) Sea ice segmentation from SAR data by convolutional transformer networks, doi: 10.48550/arXiv.2306.07649

Ronkainen I (2013) Long-term changes in Baltic Sea ice conditions. M.Sc. Thesis, Department of Physics, University of Helsinki, Finland.

Ronneberger O, Fischer P and Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, and Frangi A (eds) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015.* Lecture Notes in Computer Science, vol. 9351. Springer, Cham, doi: 10.1007/978-3-319-24574-4_28

Shelhamer E, Long J and Darrell T (2017) Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **39**(4), 640–651. doi: 10.1109/TPAMI.2016.2572683

Sinaga KP and Yang M-. S (2020) Unsupervised K-means clustering algorithm. *IEEE Access* **8**, 80716–80727. doi: 10.1109/ACCESS.2020.2988796

Stokholm A and 5 others (2022) AI4SeaIce: toward solving ambiguous SAR textures in convolutional neural networks for automatic sea ice concentration charting. *IEEE Transactions on Geoscience and Remote Sensing* **60**, 1–13. doi: 10.1109/TGRS.2022.3149323

Vaswani A and 8 others (2017) Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*. Curran Associates Inc., Red Hook, NY, USA, pp. 6000–6010, doi: 10.48550/arXiv.1706.03762

Wan H and 7 others (2023) Multi-featured sea ice classification with SAR image based on convolutional neural network. *Remote Sensing* **15**, 4014. doi: 10.3390/rs15164014

Wang YR and Li XM (2021) Arctic sea ice cover data from spaceborne synthetic aperture radar by deep learning. *Earth System Science Data* **13**(6), 2723–2742. doi: 10.5194/essd-13-2723-2021

Wang S, Chen W, Xie SM, Azzari G and Lobell DB (2020) Weakly supervised deep learning for segmentation of remote sensing imagery. *Remote Sensing* **12**(2), 207. doi: 10.3390/rs12020207

Wessel P and Smith WHF (1996) A global self-consistent, hierarchical, high-resolution shoreline database. *Journal of Geophysical Research* **101**, 8741–8743.

Yu Y and 7 others (2023) Techniques and challenges of image segmentation: a review. *Electronics* **12**(5), 1199. doi: 10.3390/electronics12051199

Yu B, Meng J, Zhang X and Ji Y (2013) Segmentation method for agglomerative hierarchical-based sea ice types polarimetric data. *Journal of Remote Sensing* **17**, 887–904.

Zhang CX and Chen SJi (2022) Semantic image segmentation for sea ice parameters recognition using deep convolutional neural networks. *International Journal of Applied Earth Observation and Geoinformation* **112**, 102885. doi: 10.1016/j.jag.2022.102885

Zhang Q and Skjetne R (2015) Image processing for identification of sea-ice flows and the flow size distributions. *IEEE Transactions on Geoscience and Remote Sensing* **53**(5), 2913–2924. doi: 10.1109/TGRS.2014.2366640

Zhang T, Ramakrishnan R and Livny M (2006) An efficient data clustering method for very large databases, SIGMOD '96: Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data, New York, NY, USA: ACM, pp. 103–114, doi: 10.1145/233269.233324

Zhu XX and 8 others (2021) Deep learning meets SAR: concepts, models, pitfalls, and perspectives. *IEEE Geoscience and Remote Sensing Magazine* **9**(4), 143–172. doi: 10.1109/MGRS.2020.3046356