

MACHINE LEARNING FOR SEA ICE MONITORING FROM SATELLITES

C.O. Dumitru^{*}, V. Andrei, G. Schwarz, M. Datcu

German Aerospace Center (DLR), Remote Sensing Technology Institute, Münchener Str. 20, 82234, Weßling, Germany
- (corneliu.dumitru, gottfried.schwarz, mihai.datcu)@dlr.de, vlad.andrei@tum.de

ICWG II/III: Pattern Analysis in Remote Sensing

KEY WORDS: Copernicus, change detection, machine learning, sea ice, maps, data analytics

ABSTRACT:

Today, radar imaging from space allows continuous and wide-area sea ice monitoring under nearly all weather conditions. To this end, we applied modern machine learning techniques to produce ice-describing semantic maps of the polar regions of the Earth. Time series of these maps can then be exploited for local and regional change maps of selected areas. What we expect, however, are fully-automated unsupervised routine classifications of sea ice regions that are needed for the rapid and reliable monitoring of shipping routes, drifting and disintegrating icebergs, snowfall and melting on ice, and other dynamic climate change indicators. Therefore, we designed and implemented an automated processing chain that analyses and interprets the specific ice-related content of high-resolution synthetic aperture radar (SAR) images. We trained this system with selected images covering various use cases allowing us to interpret these images with modern machine learning approaches. In the following, we describe a system comprising representation learning, variational inference, and auto-encoders. Test runs have already demonstrated its usefulness and stability that can pave the way towards future artificial intelligence systems extending, for instance, the current capabilities of traditional image analysis by including content-related image understanding.

1. INTRODUCTION

Currently, the European Copernicus mission with its Sentinel satellites and the free access to all their data is opening the way towards systematic large-scale scientific Earth observation data analysis relying on quality-controlled and calibrated data. In the following, we concentrate on the interpretation of satellite images delivered by Sentinel-1 (an all-weather twin SAR satellite constellation delivering polarized radar backscatter images of the Earth's surface). Due to their frequent overpasses, these satellites allow us to monitor regularly and over long time spans, for instance, the northern polar regions of our planet. Thus, the radar images delivered by Sentinel-1 (ESA S-1, 2019) can be used as a backbone for the analysis and interpretation of sea ice in the arctic and northern waters.

Existing ice classifications and frequently updated ice charts have already reached a well-defined basis for automated data analytics. Typical publicly available products being already available are sea ice index applications and medium-resolution ice charts often derived from microwave instrument data, typically providing surveys with a pixel size of about 20×20 kilometres. When we compare this resolution with what can be obtained from Sentinel images with a typical pixel spacing of 10 to 40 meters, one can immediately understand that an analysis of these high-resolution data cannot concentrate only on the details of a few selected image patches but has to include compact statistical metrics describing large ice-covered areas in global way. In particular, when we are interested in time series of images, we have to resort to tools that reduce the huge volume of pixel data to manageable quantities of compact descriptors extracted by automated tools. Then one can try to learn more about the dynamic characteristics and the behaviour of ice cover that are needed for current and future climate research. This scientific approach has to discover synergies between semantics (*i.e.*, image understanding based on

learning), and geometry (*e.g.*, automatic object extraction based on pixel neighbourhood statistics).

For an in-depth analysis of actual ice cover parameters, the Sentinel data had to be embedded in a big data processing environment (framework) that allowed efficient and reliable image analysis. This big data environment was also a precondition to fully exploit the image information content by modern machine learning methods such as deep learning and artificial intelligence as these disciplines hinge on extensive training with representative data. In our case, we were also confronted with diverse user groups, existing and future applications, and their specific requirements ranging from routine monitoring of shipping routes, and the motion of icebergs, up to the detection of climate change indicators by determining the actual ice types and their dynamic changes versus time. This resulted in extensive lists of requirements and user wishes that continuously underwent many changes. Thus, any new research should also profit from recent and ongoing developments in the field of machine learning, where we encounter rapid progress in a number of already established subsections.

These considerations led to the conception of a systematic approach for the analysis and interpretation of high-resolution ice cover images from satellites. In a first step, we compiled typical examples of sea ice cover, its life cycle and visibility over several months, and then compared existing ice cover data products and their characteristics with existing conventional image analysis tools, and our expectations from Sentinel data. This step included a survey of new technical observing opportunities offered by the Sentinel satellites as well as consultation meetings with experts in the field of ice cover analysis to learn more about their current and future needs. Finally, this first step was complemented by a list of open issues that hopefully can be solved by new approaches (ExtremeEarth, 2019).

^{*}Corresponding author

Then our second step consisted of a compilation of already existing data analytics approaches and image understanding tools from different disciplines that could be transformed into ice analysis routines and tools. In particular, this step was crucial for getting a comprehensive survey of various potential machine learning approaches and their environment (such as the selection of most appropriate reference data or useful quality criteria being needed for comprehensive testing and validation).

Our third step was then the design, implementation, test, and validation of a set of selected core routines that were tailored to the user needs. At the end, a successful implementation could be demonstrated by proving how well the combination of observing modes, instrument data, image processing, and deep analysis tools fitted with theoretical weather and ice cover modelling that was available as background knowledge from a number of external partner institutes. As a matter of fact, this had to be verified and demonstrated by examples that represent several well-understood development cycles of ice cover dynamics.

The inclusion of experiences from satellite image processing expertise combined with innovative machine learning algorithms and sea ice models did lead to a new toolset that supports the advanced analysis of ice cover phenomena. Current investigations by other researchers typically aim at detailed physical models, for instance, how ice floes disintegrate and how their backscatter changes due to weather conditions (*e.g.*, due to melting ice, fresh snowfall or wave patterns on water surfaces). When we manage to integrate all this information with optimal observing parameters of modern remote sensing instruments (*e.g.*, polarized sensing or the choice of viewing and backscatter angles), we obtain a new perspective for data interpretation.

The organisation of this contribution is as follows. Section 2 presents the characteristics of our data set. Section 3 outlines our image analysis methodology used to generate reference ground-truth data and visual-statistical analytics. Section 4 describes the image classification, semantic annotations, and change maps results using the methodology presented in Section 3. Finally, Section 5 contains conclusions and future works. The paper ends with acknowledgments and a list of references.

2. DATA SET

The selected area of interest for our use case in (ExtremeEarth, 2019) is Belgica Bank in Greenland. By accessing the Sentinels ESA hub (Copernicus, 2019) in order to select and download the images of interest, we noticed that for this area only Sentinel-1 images were available.

The Sentinel-1 data set comes from a C-band Synthetic Aperture Radar (SAR) instrument and includes a collection of five images acquired between April and December 2018 (*April 17th, June 16th, August 9th, October 10th, and December 1st*).

Figure 1 (top) shows the area of interest, while the bottom part of Figure 1 presents two out of five image quick-looks (data acquired on *April 17th, 2018* and *June 16th, 2018*).

The characteristics of the Sentinel-1 data are described in detail in (ESA S-1, 2019).

From the available Sentinel-1 data products, we selected for demonstration, based on our previous experience (Dumitru, et al., 2018), level-1 Ground Range Detected data with high

resolution taken in Interferometric Wide swath mode. The products are geo-coded with a resolution of 20×22 meters (range × azimuth) and a pixel spacing of 10×10 meters. For these products, the images are provided in dual polarization (for our polar areas, HH and HV) and with an incidence angle of about 45°. The average size of the images is 25,670×16,640 pixels.

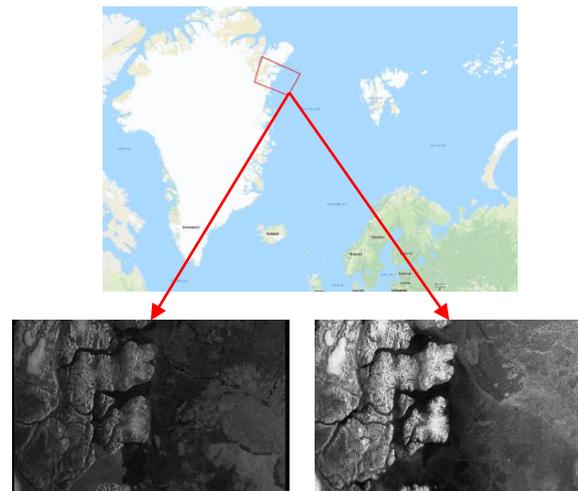


Figure 1. The location of the area of interest (top), the quick-look of the Sentinel-1 data acquired on April 17th, 2018 (bottom-left), and the quick-look of the Sentinel-1 data acquired on June 16th, 2018 (bottom-right).

3. METHODOLOGY

3.1 Semantic Annotation and Statistical Analytics Based on Active Learning

The objective of this first part of our methodology is the extraction of meaningful information (*i.e.*, knowledge) that characterizes the Earth's surface, and to semantically annotate this content. After that, based on the information extracted as semantic labels, we generated corresponding maps with semantic meaning and statistical results.

In order to attain the above-mentioned objectives, we exploited a data mining tool that is based on a Support Vector machine (SVM) and which had been developed by us in an ESA-funded project (EOLib, 2018).

This process could be divided into two parts: the first part is taking the satellite image product files and is extracting the relevant metadata and the optimal primitive descriptors that characterize the image data. For each satellite image, this information is extracted from image patches (*e.g.*, 128×128 pixels, 64×64 pixels) and ingested into a data base.

The second part is an interactive process based on a cascaded active learning method (Blanchart, et al., 2014). With the help of image analysts, the extracted information is converted into semantic descriptors that are attached to each patch. Once a full image has been annotated, different statistical analytics and/or change maps can be generated from the data base.

Figure 2 shows the two parts of the processing flow, while Figure 3 gives three examples of patches per category, with a size of 128×128 pixels, which are extracted from one of the Sentinel-1 images.

An overview of the proposed method of this section is given in the following:

Step 1 Data Pre-processing: Select the Sentinel-1 images to be processed and tile them into patches (e.g., 256×256 pixels, 128×128 pixels, etc.).

Step 2 Feature Extraction and Classification: Extract a 60-dimensional feature vector from each patch using Gabor filters (Manjunath and Ma, 1996), (Mpeg7, 2019) with five scales and six orientations (by computing the mean and variance of the patch coefficients). As an alternative, we can extract a 144-dimensional feature vector from each patch using Weber local descriptors (Chen et al., 2010) with eight orientations and 18 excitation levels.

Then we classify the feature vectors into categories using a cascaded active learning method (Blanchart, et al., 2014). Each patch is assigned to a single category based on the dominant content of the patch.

Step 3 Semantic Annotation: We annotate each category by giving it an appropriate semantic meaning (by choosing one of the categories from (Dumitru, et al., 2016)). These data can be used as reference ground-truth data set (cf. Section 3.2).

Step 4 Semantic Maps and Statistical Analytics: We generate semantic classification maps based on the annotated data from the previous step as cartographic representations and create analytics to see the distribution or the changes in an image as quantitative analyses.

3.2 Representation Learning with Variational Auto-Encoders

Representation Learning is a relatively new field that deals with finding representations of data that are needed when compiling predictors, such as classifiers and regressors (Bengio, 2013). One of the thoughts behind this paradigm is to alleviate the difficulties that occur when selecting hand-engineered features, mainly due to the fact that they do not capture the discriminative nature of the data. Because the amount of SAR imagery has increased continuously in the last years, one important task to be solved is classification. This task not only arises when one is performing semantic annotation or land-

cover classification, but also when we try to detect changes in SAR images. In the following, we devise a methodology for feature extraction and classification. This is then showcased for the practical case of change detection. At first we train a variational auto-encoder neural network (Kingma, 2014) and use the learned representations to design a feature descriptor. Afterwards, the features are fed to a robust classifier such as the *k*-Nearest Neighbours (*k*-NN) algorithm and a Support Vector Machine (SVM). The evaluation of the classification results is made with the help of precision, recall, accuracy, and F-1 score. The section is structured as follows: At first we present the theoretical background, then we propose a methodology for feature extraction and classification. At last, we present the classification results for an annotated image, and the outcome of change detection.

3.2.1 Variational Inference and Auto-Encoders

A common problem in Bayesian statistics is computing the likelihood $p(x)$ of the data. As this can only be exactly computed in certain special cases, one tries to approximate it with a distribution $p_\theta(x)$, by maximizing the log-likelihood under the parameters θ . The data, e.g., 128×128×2 patches of Sentinel-1 SAR imagery, additionally depend on some hidden variables z . As also the true posterior $p_\theta(z|x)$ is in general untractable, a lower bound is maximized instead. Following Bayes' rule, a marginalization over the hidden variables, and Jensen's inequality (Beal, 2003), such a lower bound can be written as:

$$\ln(p_\theta(x)) \geq \mathbb{E}_{q_\varphi(z|x)}[\ln(p_\theta(x|z))] - KL(q_\varphi(z|x) || p_\theta(z)) = \mathcal{L}(x, \varphi, \theta), \quad (1)$$

where $KL(\cdot || \cdot)$ is the Kullback-Leibler divergence (Cover and Thomas, 2012). The main idea in variational Bayes' theory is to learn the set of parameters $(\varphi, \theta)_{opt}$ which maximize $\mathcal{L}(x, \varphi, \theta)$.

If one uses neural networks to estimate the parameters for both models (Kingma, 2017), then one talks about a variational auto-encoder. In the following, we suppose the hidden variables $p_\theta(z)$ come from an isotropic Gaussian $\mathcal{N}(0, I)$ and the encoder is also a Gaussian with a mean vector μ and a diagonal covariance matrix $\sigma^2 I$.

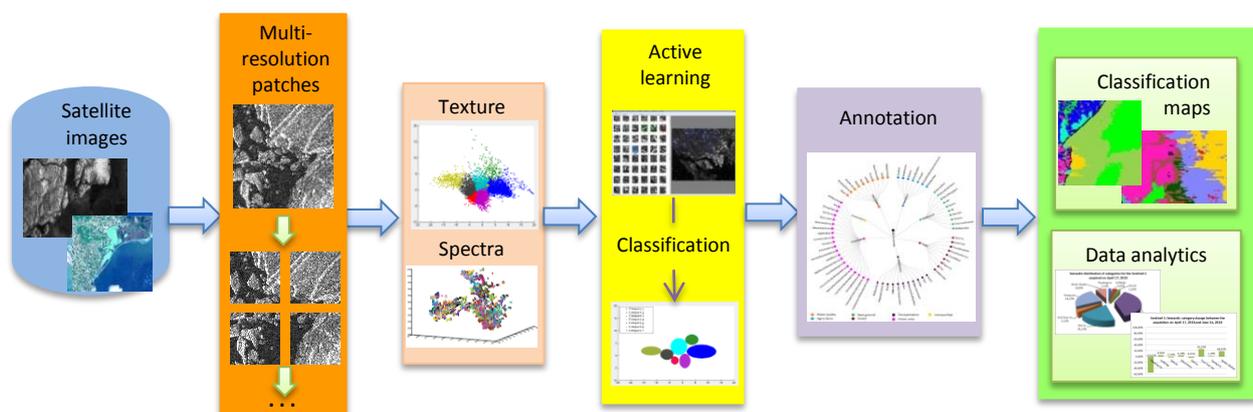


Figure 2. Processing flow used to generate semantically annotated data, classification maps, and analytics.

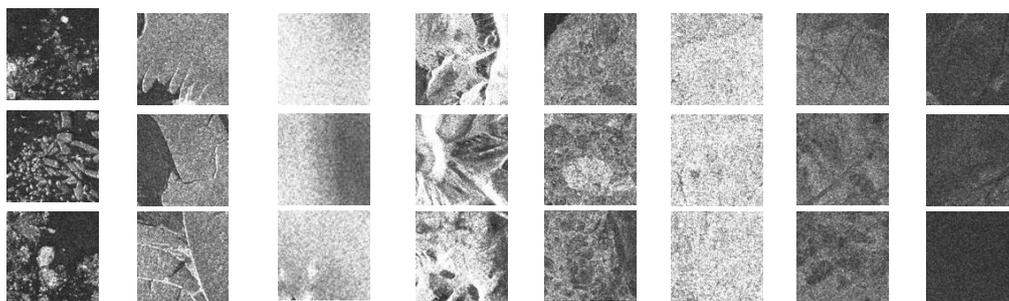


Figure 3. Examples of three classified patches per column: The semantic meaning of each category (from left to right) is: *Floating Ice, Icebergs, Glaciers, Mountains, Old Ice, First-Year Ice, Young Ice, and Water Bodies.*

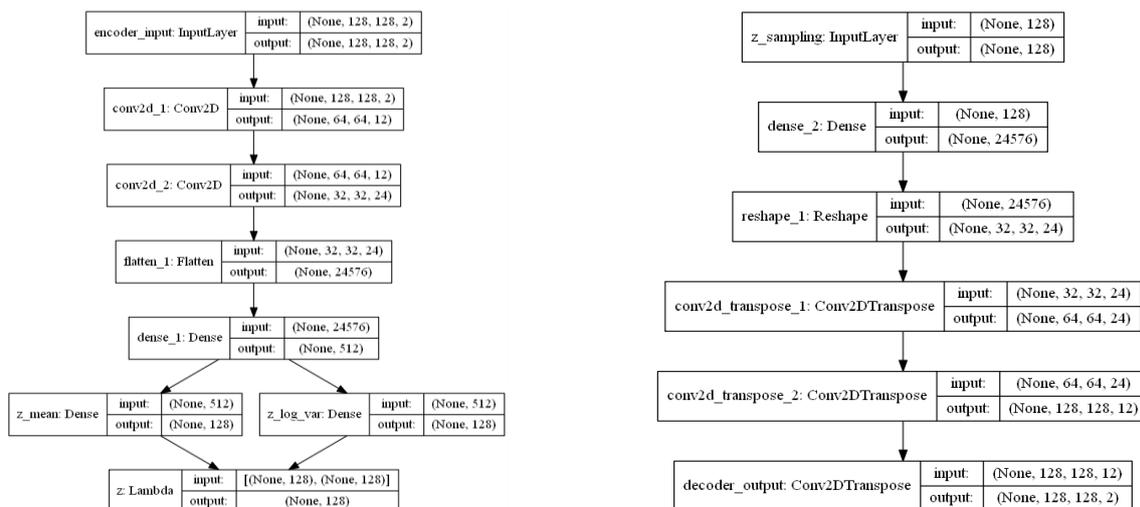


Figure 4 (left). Network structure of the encoder. “None” refers to the batch size being used for training which, in this case, was 32. (right) Network structure of the decoder.



Figure 5.a. Sentinel-1 quick-look view (left) and a classification map (right) for an image of Belgica Bank, Greenland acquired on April 17th, 2018.

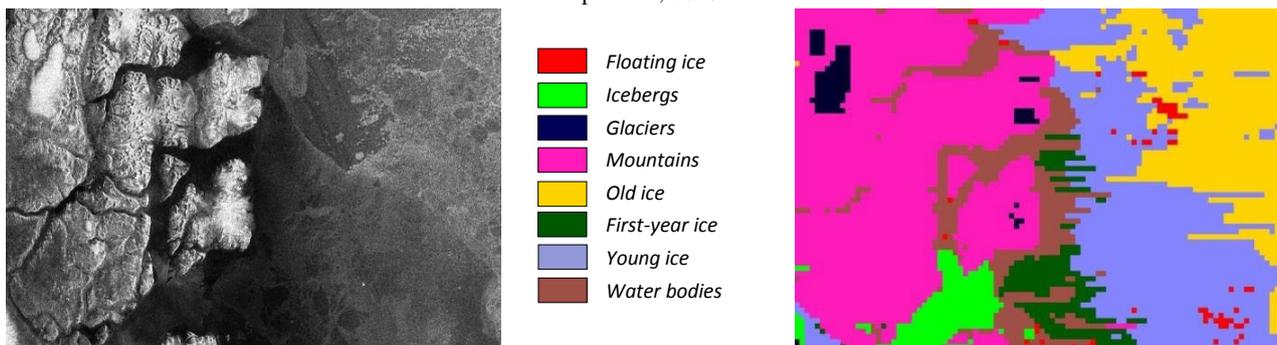


Figure 5.b. Sentinel-1 quick-look view (left) and a classification map (right) for an image of Belgica Bank, Greenland acquired on June 16th, 2018.

An overview of the proposed method is given by:

Step 1 Data Pre-processing: Eleven Sentinel-1 images were downloaded, covering all kinds of terrain and water for training the network. The images were tiled into 128×128×2-sized patches, and from all patches 250,000 of them were used for training, and 13,300 for validation.

Step 2: Network Training: A network architecture was designed and trained using the ADAM optimizer (Kingma, 2015), until the lower bound reached its minimum, in our case 0.359. The architecture for the encoder and decoder networks is illustrated in Figure 4 (left) and Figure 4 (right). The encoder consists of a convolutional front-end, made up of two layers, and several fully connected layers used to estimate the parameters μ and σ . The decoder network is designed symmetrically, using deconvolutional layers.

Step 3: Feature Extraction and Classification: The image whose features need to be extracted is convolved with the filters of the encoder’s convolutional layers. From each image the mean and variance are computed, and the results are concatenated to a single 72-dimensional feature vector φ , which is then fed to the aforementioned classifiers. For demonstration, we are using: a) Support Vector Machines (Chand and Lin, 2001) which are powerful classifiers that have been used in remote sensing applications for many years (Mountrakis et al, 2011). The best hyperparameters C and γ were found to be $C = 1000$ and $\gamma = 0.001$ by cross-validation. b) Another classifier is the k -Nearest Neighbours algorithm. The choice of a metric as well as the constant k is up to the user. In our experiments, we found out that weighted k -NN works best with a Euclidean metric and $k = 9$.

Step 4: Change Detection: The detection of changes was performed for all images of the data set, taking the image acquired on April 17th, 2018 as a reference. Because a binary change detection would not give so much insight into the nature of the changes, we devised the following strategy to quantify a change. We took the predicted labels for each patch from the images, compared them with the label in the reference image, and then computed the difference counts. Based on this, we were defining different levels of change. Mathematically, this can be summarized by following formula:

$$ChangeLevel = |y_{ref} - y_{pred}|, \quad (2)$$

where y_{ref} denotes the label of an image patch found in the reference image and y_{pred} is the predicted label of the unannotated patch. The labels are integers and correspond to the classes in the annotation, as follows:

- 1 – Floating Ice
- 2 – Icebergs
- 3 - Glaciers
- 4 – Mountains
- 5 – Old Ice
- 6 – First-Year Ice
- 7 – Young Ice
- 8 – Water Bodies

For example, if an image patch of *Floating Ice* in the reference data changes into a *Water Bodies* patch, the absolute value of the difference between the labels, *i.e.* the change level will be 7.

Note that the change levels are bound to the ordering of the categories/classes and could be defined differently.

4. EXPERIMENTAL RESULTS

4.1 Reference Data Set, Classification Maps, and Statistical Analytics

By using the method described in Section 3.1, we are able to generate a reference ground-truth data set for our area of interest that was used later in Section 3.2.

For demonstration, we chose two out of five images, acquired on April 17th, 2018 and June 16th, 2018. Figure 5.a and Figure 5.b show the results of the semantic classification and the diversity of categories that the method is able to retrieve from the images.

Figure 6 (top and bottom) displays the statistical analytics of our method that was able to retrieve and semantically annotate eight categories using (Dumitru, et al., 2016). By analysing in detail the diversity of each semantic category, we could observe the changes that occurred within two months in the area of interest. Similar changes occurred when we compared the other images. Figure 6 (center) presents the differences (increase or decrease) of the semantic categories between the inspected image pairs.

4.2 Classification Results of the Annotated Data Set and Subsequent Change Detection

Table 1 shows the classification results of the annotated data sets using the method from Section 3.2. The presented metrics are averaged over all eight semantic categories in the data set. The classifiers were trained on 20% of the labelled data and tested on the rest.

The k -NN classifier has a robust performance for both images (the ones acquired on April 17th and June 16th, 2018), the classification metrics was only slightly decreasing when classification was performed on the second one. For SVM, the classification results show a bigger overall drop. That is why, in the following, we decided to use k -NN for the change detection task.

In the following, we present the change detection results. This was done with the image of April 14th as a reference (see Figure 5.a right side), while the images dated June 16th, August 9th, October 14th, and December 12th, 2018 were compared with it. After **step 4**, we could find eight levels of change. Note that the transitions might work in both directions, as we used the absolute value of the difference (see Table 2).

Figures 7 to 10 show the change maps for different images.

| Classifier | Precision | Recall | F-1 | Accuracy |
|-----------------------------------|-----------|--------|-----|----------|
| k -NN April 17 th | 89% | 89% | 88% | 89% |
| SVM April 17 th | 90% | 88% | 87% | 88% |
| k -NN June 16 th | 87% | 88% | 87% | 88% |
| SVM June 16 th | 84% | 82% | 83% | 82% |

Table 1. The average performances obtained over all categories using k -NN and SVM for two acquisitions.

We could see that the change maps portray all kinds of possible transitions that are caused by the changing seasons. Additional expert validation would be helpful in order to understand the model correctness and accuracy.

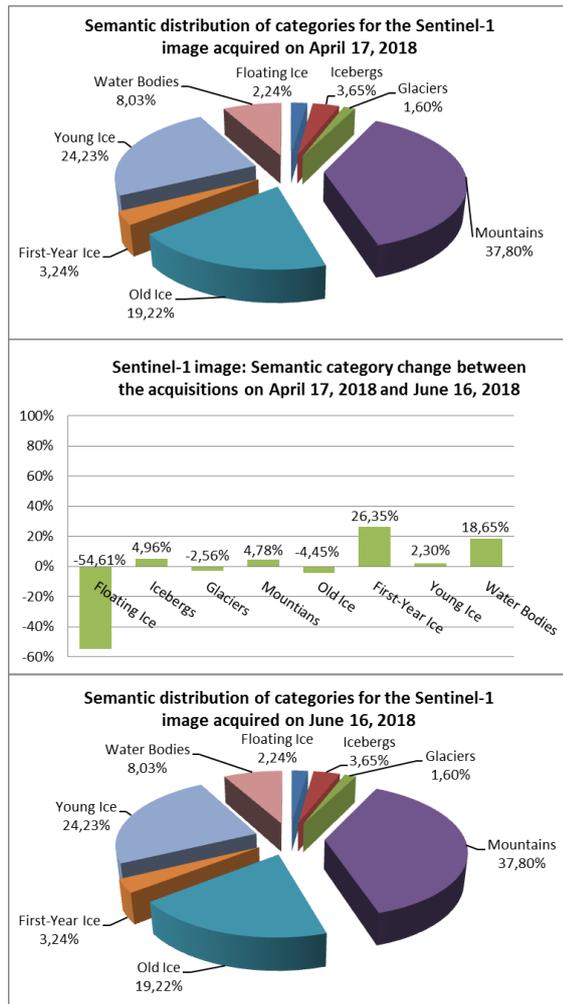


Figure 6. Quantitative analysis of the semantic categories that were extracted from the area of interest using the Sentinel-1 image acquired on April 17th, 2018 (top) and the Sentinel-1 image acquired on June 16th, 2018 (bottom). The semantic change was determined also quantitatively (center).

| Levels | Transition between categories/labels |
|----------------|---|
| <i>level 0</i> | No change |
| <i>level 1</i> | Old Ice – First-Year Ice, First-Year Ice – Young Ice Young Ice – Water Bodies |
| <i>level 2</i> | Floating Ice – Glaciers, Glaciers – Old ice, Old Ice – Young Ice First-Year Ice – Water Bodies |
| <i>level 3</i> | Icebergs – Old Ice, Old Ice – Water Bodies |
| <i>level 4</i> | Floating Ice – Old Ice |
| <i>level 5</i> | Floating Ice – First-Year Ice, Icebergs – Young Ice |
| <i>level 6</i> | Floating Ice – Young Ice Icebergs – Water Bodies |
| <i>level 7</i> | Floating Ice – Water bodies |

Table 2. Different levels of change.

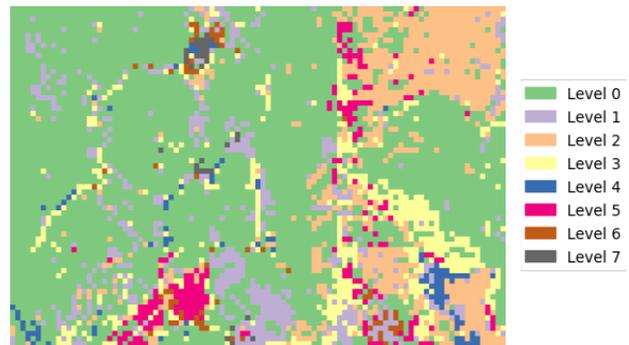


Figure 7. Changes between April 17th and June 16th, 2018.

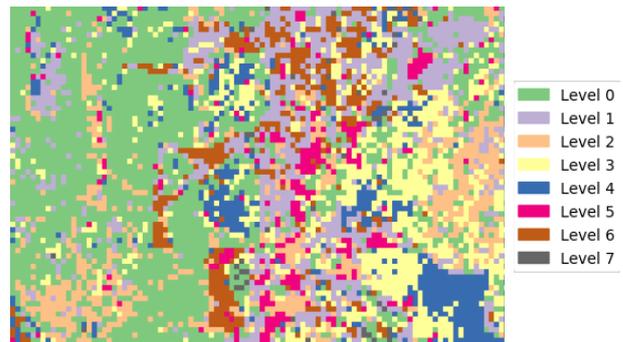


Figure 8. Changes between April 17th and August 8th, 2018

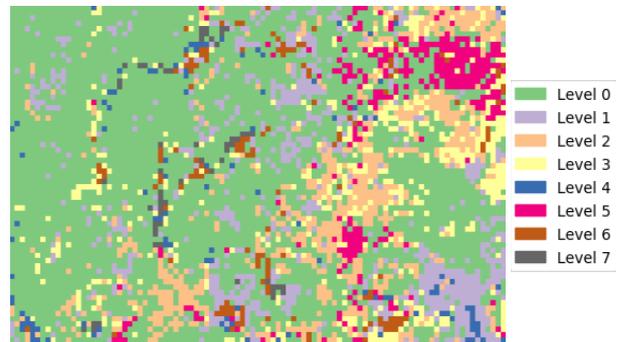


Figure 9. Changes between April 17th and October 14th, 2018.

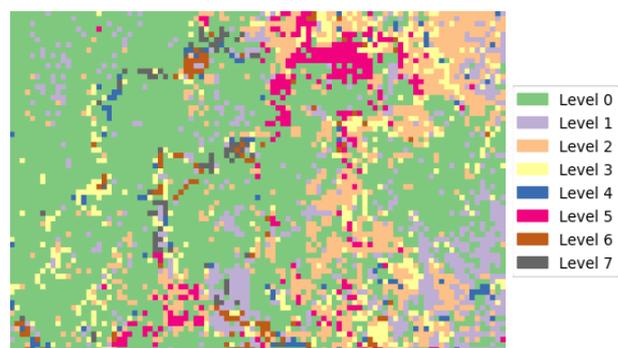


Figure 10. Changes between April 17th and December 12th, 2018.

5. CONCLUSIONS AND FUTURE WORKS

In summary, this paper presents three points:

- a method of active learning for the generation of reference data and analytics;

- a method of learning representations of Sentinel-1 SAR data; and
- a way of quantifying changes, by computing the absolute values of the differences between the semantic categories/labels.

It would also be interesting to explore other ways of enhancing the change detections, by learning the physical parameters of the objects with a neural network, and using these parameters as features, or by devising more levels of change, which can be done if one uses the positive/negative differences between the labels instead of the absolute value of them. All these are a good start for future work.

During the period of the period (ExtremeEarth, 2019) more areas will be investigated by polar ice experts for still more detailed validation of the proposed methods.

The future use of automated SAR data interpretation will depend on the easy and timely access to high-quality image data and their ease of handling. We expect that the capabilities of the current Sentinel-1 instruments give the user community already a good chance to get acquainted with advanced machine learning approaches applied to SAR images – more to come during the next years!

ACKNOWLEDGEMENTS

We will like to thank Nick Hughes from the Norwegian Meteorological Institute in Oslo, Norway (a partner in the H2020 ExtremeEarth project), for helpful discussions about the investigated area.

REFERENCES

Beal, M. J., 2003. Variational Algorithms for Approximate Bayesian Inference, *PhD Dissertation*, University College London.

Bengio, Y., Courville, A.C., Vincent, P., 2013. Representation Learning: A Review and New Perspectives, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 35(8), 1798-1828.

Blanchart, P., Ferecatu, M., Cui, S., Datcu, M., 2014. Pattern Retrieval in Large Image Databases Using Multiscale Coarse-to-Fine Cascaded Active Learning, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(4), 1127-1141.

Chen, J., Shan, S., He, C., Zhao, G., Pietikainen, M., Chen, X., Gao, W., 2010. WLD: A robust local image descriptor, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9), 1705–1720.

Chang, C.C., Lin, C.J., 2001. LIBSVM: a Library for Support Vector Machines. Available online <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>.

Copernicus Open Access Hub, June 2019. Available online: <https://scihub.copernicus.eu/dhus/>.

Cover, T.M., Thomas, J.A., 2012. Elements of Information Theory, 2nd edition, John Wiley & Sons, ISBN 1118585771.

Dumitru, C.O., Schwarz, G., Datcu, M., 2016. Land Cover Semantic Annotation Derived from High-Resolution SAR Images, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(6), 2215-2232.

Dumitru, C.O., Schwarz, G., Datcu, M., 2018. SAR Image Land Cover Datasets for Classification Benchmarking of Temporal Changes, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(5), 1571-1592.

EOLib project, February 2018. Available online: <http://wiki.services.eoportal.org/tiki-index.php?page=EOLib>.

ESA Sentinel-1 portal, May 2019. Available online: <https://sentinel.esa.int/web/sentinel/missions/sentinel-1/satellite-description/geographical-coverage>.

ExtremeEarth project, July 2019. Available online: <http://earthanalytics.eu/>.

Kingma, D.P., Ba, J., 2015. Adam: A Method for Stochastic Optimization, *Proc. of the 3rd International Conference on Learning Representations*, arXiv: 1412.6980v9.

Kingma, D.P., Welling, M., 2014. Auto-Encoding Variational Bayes, *Proc. of the 2nd International Conference on Learning Representations*, arXiv: 1312.6114.

Kingma, D.P., 2017. Deep Learning and Variational Inference: A New Synthesis. Available online: <https://www.uva.nl/binaries/content/assets/uva/nl/onderzoek/pro moveren/samenvattingen/2017/10/samenvatting-kingma.pdf>.

Manjunath, B.S., Ma, W.Y., 1996. Texture features for browsing and retrieval of image data, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8), 837-842.

Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review, *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247-259.

MPEG7 standard, 2019. Available online: <http://mpeg.chiariglione.org/standards/>.