POAC'19

Delft, The Netherlands

Proceedings of the 25th International Conference on
Port and Ocean Engineering under Arctic Conditions
June 09-13, 2019, Delft, The Netherlands

Semi-Automatic Ice Floe Detection for Drift Evaluation

Yevgeniy Kadranov[1], Sergey Vernyayev[1], Anton Sigitov[1]

[1] LLP ICEMAN.KZ, Shymkent, Kazakhstan

ABSTRACT

Automatic ice drift detection from remote sensing data is a straight forward process in the arctic where revisit time by polar orbiting satellites is quite small and it is possible to rely on one satellite to define change detection algorithms and extract drift data. But if we consider sub-Arctic regions intervals of about three days minimum between images of the same satellite makes automated drift detection obsolete for majority of cases due to changes in wind regime and intermittent drift events. Therefore, multi-platform remote sensing data needs to be used to reduce the gaps. In this project we exploit deep learning capabilities in order to estimate ice drift between image from different sensors, by learning their similarities with Siamese Neural Network.

KEY WORDS: Ice Floe Recognition; Ice Drift detection; Remote Sensing; Ice Charting Operations, Deep Learning, Computer Vision

INTORODUCTION

One of the common tasks in sea ice monitoring operations is to measure ice drift from satellite images. Although it is not as accurate as drift buoy data, for example, it gives information on ice drift distribution over vast areas. ICEMAN.KZ has developed and algorithm of detecting unique floes and tracking their displacement through the season (Kadranov et al, 2017) and implemented in their ice charting processes both for drift evaluation at current time and for forecasting in the near future based on forecast model data. This study is targeted to improve performance and timeliness of previously implemented techniques by automating floe detection using multiple satellite platforms thus increasing number of observations and reducing human interpretation bias.

Many research studies were conducted in order to automatically estimate ice motion from two consecutive images, for example, by computing and matching key points (Muckenhuber 2016). These studies were mostly conducted between images of the same platform which makes sense in the Arctic as majority of satellites being polar orbiting revisit the same area at relatively short intervals allowing to assess drift correctly. Comparison of images from different sensors is less common, but, nonetheless, is important for sub-Arctic areas such as the Northern Caspian Sea being located between latitudes of 47°N and 44°N, where revisit of the same satellite can be up to 4-5 days. With such intervals between images detected changes become obsolete due to changes in wind regime and intermittent drift events.

This project is targeted to explore capabilities of deep learning to detect similar objects in consequent temporally images that were acquired with different sensors. Convolutional Neural

Networks (CNN) to find corresponding areas between Synthetic Aperture Radar (SAR) sensor data -Sentinel-1and multispectral Sentinel-2 images were used to assess applicability of the approach before introducing it into routine operations.

CONVOLUTIONAL NEURAL NETWORKS BACKGROUND

CNN show high performance on the vision-based tasks (such as classification, detection, segmentation, etc.). Classic CNN consist of input layer, convolutional layers, fully connected layers, loss function and output. Weights of the network are optimized with backpropagation – error (or loss) between output of the network is iteratively compared with expected output (label) by computing gradient of the loss function. The input data is fed to the network in small parts (batches). One iteration when full dataset passed forward and backwards the network is called Epoch. More detailed description of CNNs is discussed by Karpathy (2018).

Detection of the same floe between 2 images from different sensors was perceived as deep metric learning problem. Similarities were estimated with Siamese Neural Network, which consist of 2 identical subnetworks that share weights and output similarity score between two inputs. Experience with deep learning algorithms usage in sea ice monitoring applications is scarcely available in public, but this architecture showed good performance in such problems as face recognition (Varior et al. 2016). Pseudo Siamese Neural Network (non-shared weights) was applied by Hughes et al. (2018) for comparison of SAR and Optical data to detect corresponding patches and suggested most optimistic way forward within this study.

Siamese Neural Network was trained by showing the network pairs of Sentinel-1 and Sentinel-2 images that cover the same ice area and expecting to output low dissimilarity and then showing pair of the same Sentinel-1 image but with Sentinel-2 image that covers different ice area and expecting to output higher dissimilarity. The network was expected to learn optimal parameters to be able correctly measure dissimilarities on the unseen samples with sufficiently high number of training samples. Since one of the main conditions for CNNs good performance is quality and quantity of the labelled data, training set of images consisted of observations within stationary ice areas. This way manual matching of the corresponding floes was avoided increasing speed of training data generation. Consequently, trained network was tested over mobile areas. Geographically samples were collected over the North-East Caspian Sea giving opportunity to verify results with vast archive of observations generated in company for internal use.

TECHNICAL SET-UP

Technically the project was set up utilizing only free open source data and software.

Vast archive of Sentinel satellite data since 2014 was downloaded from Scientific hub and ensured enough samples both for training of networks and for further tests of applicability in operations. VV polarization data from Sentinel-1 acquired in IW mode (the only mode available for the Caspian region) with GRD processing and only Near Infrared band for Sentinel-2 were used to maintain simplicity of interpretation and reduce amount of image processing in the scope of this project. The images were scaled to 8 bit depth.

Images were post-processed with GDAL tools using Python scripts developed internally to adjust imagery for ice charting processes with no human interference during operations. QGIS was used for GIS based processing such as reprojection and visualization of intermediate results with algorithms clearly described by QGIS Development Team (2009) in software documentation.

CNN was trained on "Google Colaboratory" notebook that utilizes Tesla K80 GPU using Keras

API for Tensorflow framework. Training time, with final architecture (see Figure 1) on the whole training set was 330 seconds per epoch. Inference with trained model was performed on the laptop CPU Intel Core M-5Y10 0.8Ghz, the comparison time was about 0.7 seconds per pair.

Siamese Neural Network

Siamese neural network consists of 2 identical CNNs that share architecture and weights. Each network is trained by interchangeably feeding it with positive (images are the same) and negative (images are different) pairs. The output of the last layers (N-dimensional vector) of 2 CNNs are sent to contrastive loss function. This function outputs higher loss if Euclidean distance between the vectors is large for positive pairs or if the distance is small for negative pairs. Therefore, network tries to learn such parameters that would minimize distance between positive pairs and maximize it between negative pairs.

The Euclidean distance between output of the last layer of 2 subnetworks is used as dissimilarity measure to asses accuracy during training and validation stage and to detect floes during testing.

Siamese Neural Network with 2 subnetworks was built for this project based on VGG architecture by Simonyan et al (2014), one the most used CNNs for image classification, that consist of 5 convolutional blocks and 2 fully connected layers (see Figure 1). The last convolutional layer was eventually removed due to the heavy overfitting resulting in lack of generalization. It gave good performance on training set, but poor results during validation.
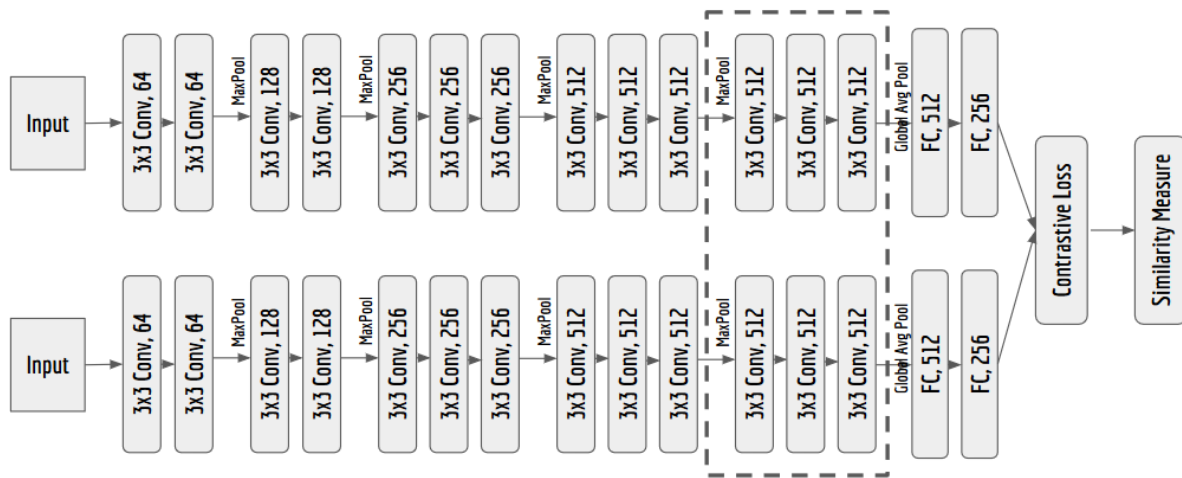

Figure 1. Architecture of Siamese Neural Network used in this project. Convolutional block that was removed in the final version of the network is bounded with dashed box.

Data Preparation

The training data was generated by splitting areas of stationary ice captured with Sentinel-1 and Sentinel-2 images on same or nearly same day into tiles with dimensions of 200x200 pixels (which spatially corresponds to 2000x2000 meters) with stride of 100 px. Figure 2 illustrates example of tile generation. The same figure shows samples of Sentinel-1 and Sentinel-2 images with the same distinctive ice floes to illustrate applicability of tile sizes compared to the size of tracked floes.

Areas of stationary ice were used to simplify generation of positive pair. Negative pairs are generated by randomly picking a tile form Sentinel-2 set (that excludes tile from reference location).

In total there were 7369 pairs of tiles generated. However, some of these tiles only included regions with homogeneous conditions, for instance, open water or flat ice without distinct floes that could be detected for analysis. Additional step of removing these areas from training and subsequently from analysis significantly reduced computation time at later stages. These tiles were removed from the set by thresholding mean squared error (MSE) between original image and the same image smoothed with gaussian filter with size of 10x10. By analyzing results, all tiles with MSE below 0.03 were considered uniform and filtered out from the set for analysis. Only 5603 pairs remained as final set for training and validation as result of filtering in such way.
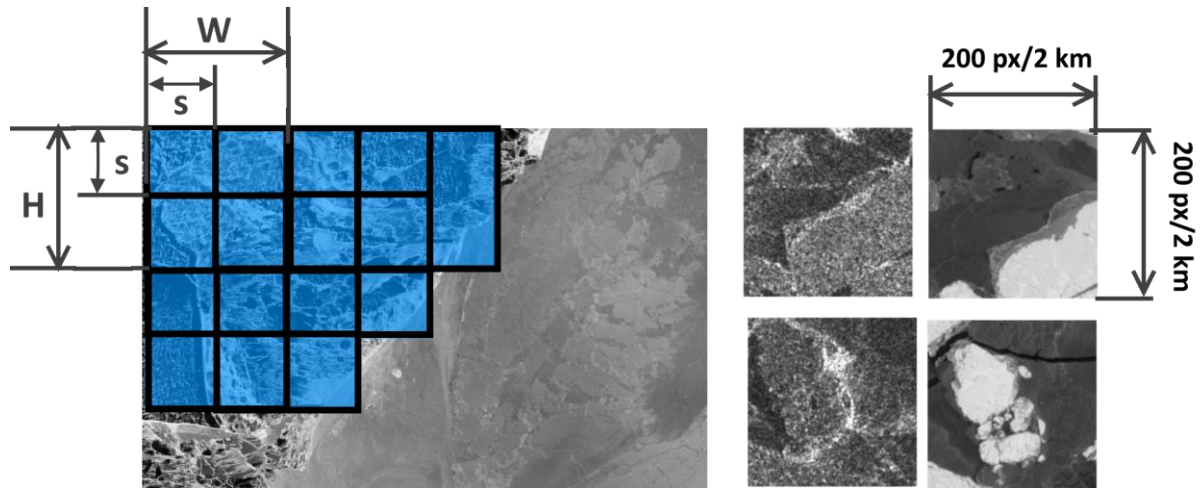

Figure 2. Left: Example of tile generation: H - Height, W - Width, s- Stride; Right - Enlarged samples of Sentinel-1 and Sentinle-2 NIR Band tiles

Training and Validation Routines

80% of tiles within the final set (4482 pairs) were randomly selected for training and the rest 20% were used for validation (1121 pairs). Tiles of the training set were used as input for the model. Validation set was used to assess performance of the model on the unseen samples.

Finally, once the architecture of the network was chosen and hyperparameters were tweaked, and, thus, the model was trained, to test the model on detection of drifting ice, another set of tiles was created from areas with mobile ice containing Sentinel-1 and Sentinel-2 which were acquired at least one day apart.

Areas with mobile floes from the test set were manually picked and then sent to the model for detection. The results were then qualitatively compared with manually detected during other projects (Kadranov et al, 2017 being part of it).

NETWORKS ASSESSMENT

Accuracy of correct classification was used for the initial evaluation based on the following logic. If the output dissimilarity was below or equal 0.5 then the pair of tiles was considered similar indicating detected floes, if above then the pair was considered different. Accuracy was calculated as ratio of correct classifications to the total number of pairs.

The first network was trained for 50 epochs with batch size of 32 pairs, making 4482/32 steps per epoch to define the most suitable architecture. Figure 3 shows accuracy ratings for networks with 5 and 4 convolutional blocks

Network with 5 convolutional blocks showed good performance on the training set, reaching almost 100% accuracy in the last epochs. However, validation accuracy stopped increasing

after 20th epoch and was around 75% for the rest of the assessment.

Networks without 5th convolutional block showed slower growth in training accuracy, but difference between validation and training accuracy was negligible. Training and validation accuracy reached 78% and 77% respectively on the last epoch.
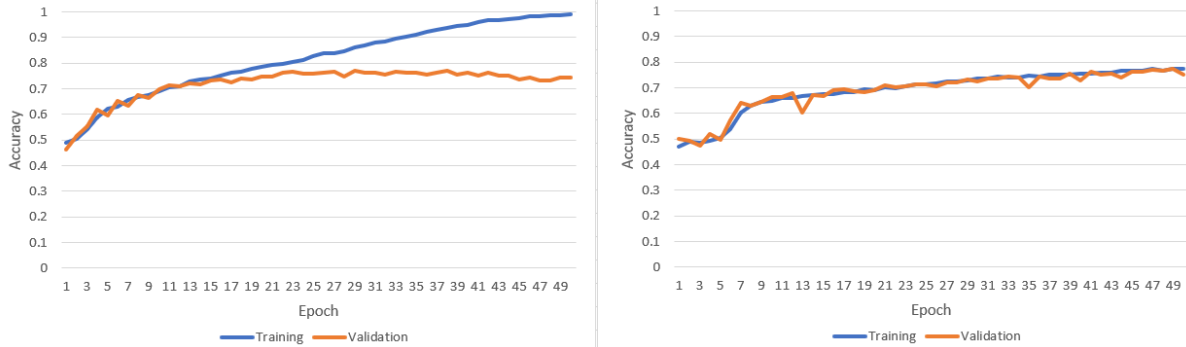

Figure 3. Left - Accuracy of the network with 5 convolutional blocks, showing overfitting; Right: Accuracy of the network with 4 convolutional blocks.

Network with 4 convolutional blocks was chosen as a final model based on the initial assessment above and was trained for 100 epochs. On the last epoch training and validation accuracies were 84% and 83% respectively as illustrated in Figure 4.
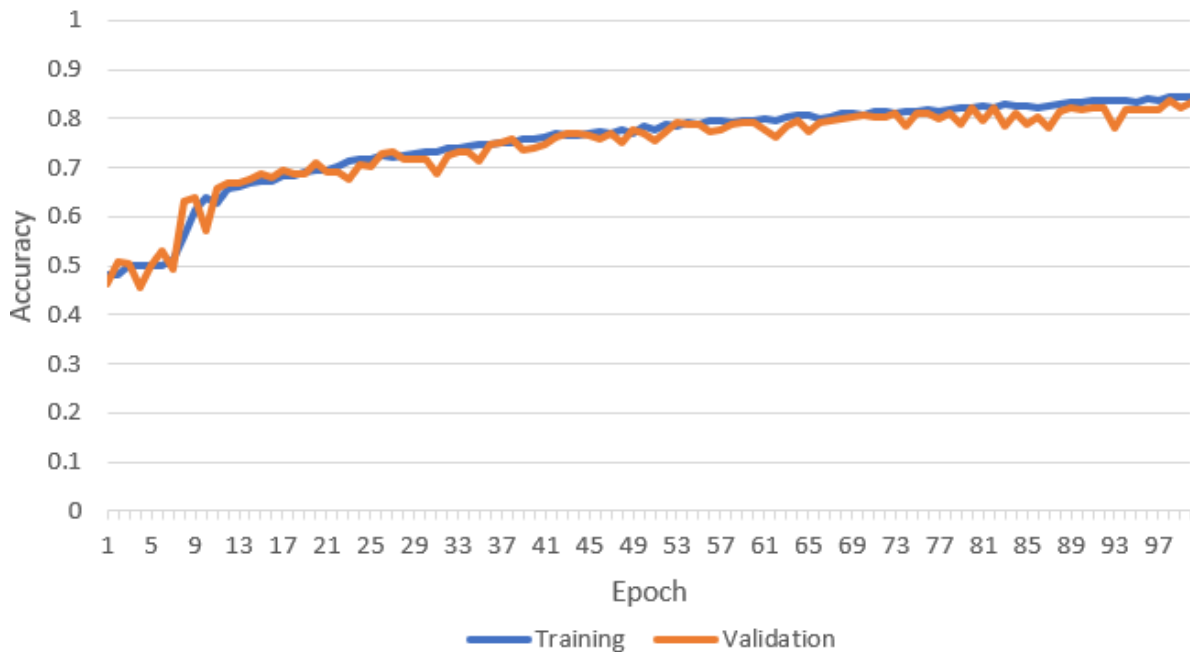

Figure 4. Accuracy of the network with 4 convolutional blocks, 100 epochs.

Training model with such accuracy was expected to detect similar floes in majority of cases between two images. However, certain number of erroneous detections can be expected after processing making it a requirement for visual check of results before further steps are taken in drift assessment. Results of successful and erroneous detections are shown below.

TRAINED MODEL PERFORMANCE TEST

Performance of the trained model on detecting the similar floe was then tested on images over mobile areas. To capture larger area tiles for this test were taken with spatial extend of 4x4 km,

making 400x400 px size and with stride of 2 km (200 px), and then resized to fit the input size of 200x200 px. Sentinel-1 was selected as reference image imitating a previous day during operations. Consequent Sentinel-2 image and corresponding areas of Sentinel-1 were then split into tiles for analysis within a search radius depending on wind conditions during the period between the two images. Each tile of Sentinel-2 as pair with reference Sentinel-1 tile was sent to Network for analysis. Pairs with minimal dissimilarity that were then chosen as a prediction of resulting detected floes would identify drifted objects for the following analysis of drift between them.

Figures below illustrate some of the resulting detections. The network has shown ability to identify similarities under varying conditions. Example illustrated in Figure 5 shows conglomerate of thicker floes drifting away from stationary ice border towards areas with smaller ice concentrations. This scenario of unconstrained drift was most optimistic as suggests minimal deformation to floes. Indeed, visually observed floes in blue box on Sentinel-2 (right) were found similar to those in red box on Sentinel-1 on the left indicating the area they originated from near stationary ice border.
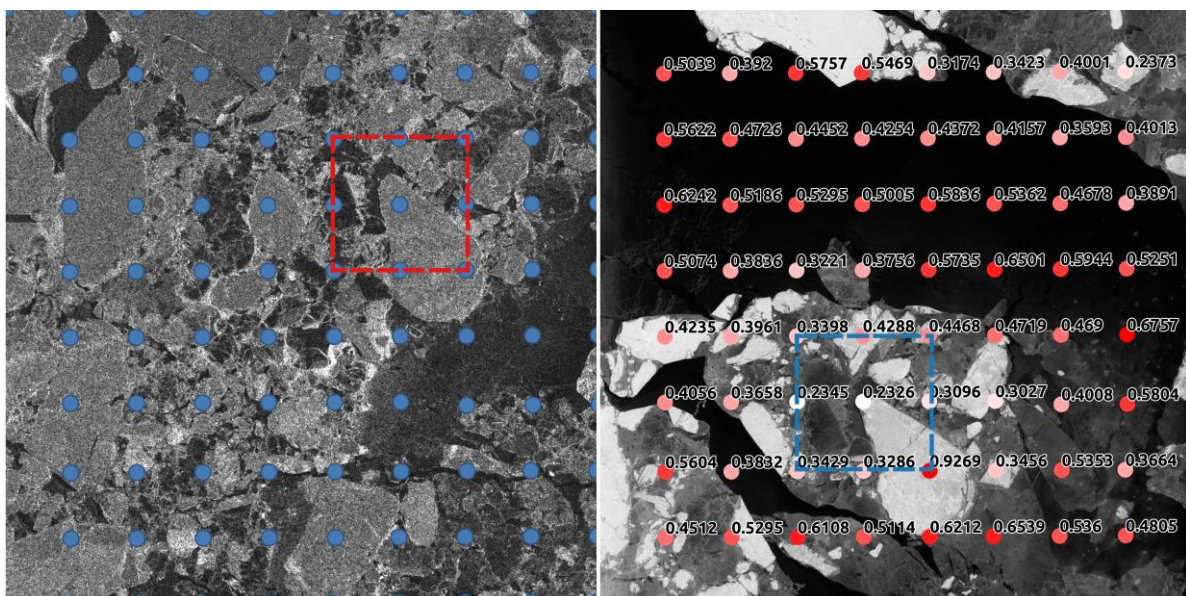

Figure 5. Example of successful floe detection (perfect match). Left - Sentinel-1 Image. blue dots are centroids of the tiles red bounding box - reference tile. Right: Sentinel-2 image acquired 1 day after the Sentinel-1, red dots –disimilarities with reference tile of Sentinel-1 measured by the network. Blue box – minimal dicimilarity from the reference tile.

Figure 6 below shows similar drift scenario but with slightly worse match of automatic identification by network to manual observation. Although the error did not exceed one tile making detection results useful for further analysis within tolerable proximity to real objects identified for consequent delineation of floes and drift vector calculations.
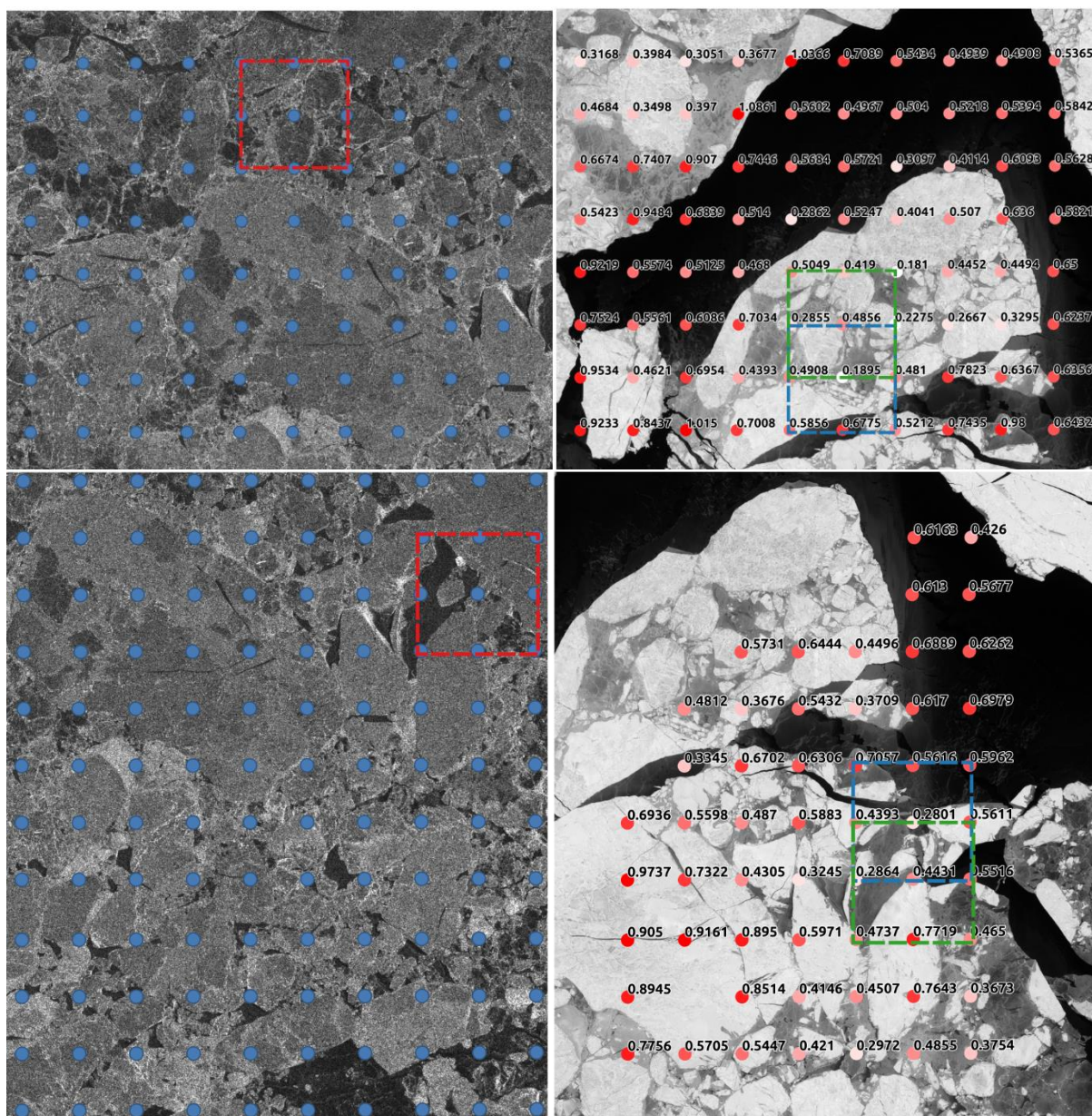
Figure 6. Example of successful floe detection (tolerable error). Left - Sentinel-1 Image. blue dots are centroids of the tiles red bounding box - reference tile. Right: Sentinel-2 image acquired 1 day after the Sentinel-1, red dots –dissimilarities with reference tile of Sentinel-1 measured by the network. Blue box – minimal disimilarity from the reference tile, green box manual detection.

In rare cases, ice floes prediction was totally incorrect, with model returning minimum dissimilarity several tiles away from the actual location of the floe. Figure 7 shows green box detected by operator and blue by network not even closely matching each other. Although these errors are spotted on the next stages of drift evaluation and do not affect results filtering them out increases processing time. Further training of model with more data fed into the network and making it more sophisticated with adding augmentation of source imagery with added iterations of processing will significantly reduce number of errors during detection.
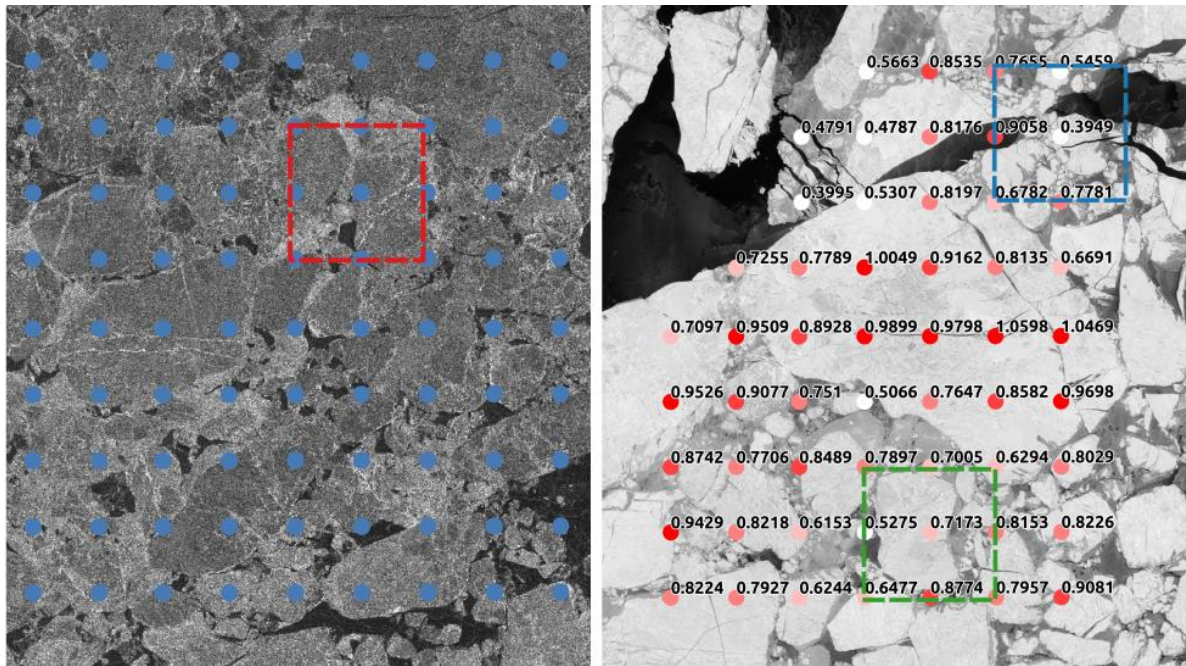
Figure 7. Example of failed prediction. Left - Sentinel-1 Image. blue dots are centroids of the tiles red bounding box - reference tile. Right: Sentinel-2 image acquired 1 day after the Sentinel-1, red dots –disimilarities with reference tile of Sentinel-1 measured by the network. Blue box – minimal dissimilarity from the reference tile, green box manual detection.

DISCUSSION

Although the detection of the ice floes was not always correct, the experiments showed that the Siamese Neural Network has high potential for multi-sensor ice drift estimation as proven with results of Sentinel images comparison for several cases in this project. Full-scale regional testing of this algorithm is yet to be conducted during further operational usage. Experience gained during this project has identified several important conclusions that will facilitate further research of fully automatic floe detection with future projects.

So, models trained on stationary ice areas has shown ability to detect floes over mobile areas. This proven ability saves significant time during model training stages by removing manual input to show similarities between images.

Demonstrated model's ability to detect floes with spatial resolution differing from the training set is quite important as well. Training and validation performed on the tiles with spatial resolution of 2x2 km took significantly longer time compared to consequent experiments with 4x4 km tile resolution. This reduction in processing time makes the whole concept of floe identification timely for operational use giving results shortly after image acquisition.

Accuracy of model at this stage still requires human interference to identify obviously erroneous detections. However, further training sessions and improvements can lead to better model performance. Additional training sets will also improve the model by achieving robustness to morphological changes of the floes occurring during displacement through areas of compaction or when splitting into smaller floes.

Current model has proven good performance on direct wind driven events of ice drift. However, changing wind direction and effects of currents during longer periods than one day lead to rotational displacements of floes and cannot be easily detected with current structure. Further improvements, however, are possible with introduction of augmented images with added

rotation, mirroring and similar distortions.

In addition to these enhancements discussed above this method of identifying floes needs more training and validation with other satellite data such as Sentinel-3, MODIS, Landsat and different SAR platforms to increase number of possible displacement observations with increasing number of cases. However, this analysis is more challenging with introduction of higher variability of input image parameters such as different spatial coverage and pixel resolution that is in most cases is not as easy to compare as with similar between Sentinel-1 and 2.

REFERENCES

ESA, 2014-2019. Copernicus open access hub. Retrieved from http://https://scihub.copernicus.eu/dhus [Accessed continuously since 2014]

Chollet F., 2015 Keras: Deep Learning for humans. Available at https://github.com/fchollet/keras. [Accessed February 1, 2019].

Kadranov, Y., Sigitov, A., Vernyayev, S, 2017. Comparison of satellite imagery based ice drift with wind model for the Caspian Sea. Port and Ocean Engineering under Arctic Conditions. Busan, South Korea.

Karpathy, A., 2018. CS231n Convolutional Neural Networks for Visual Recognition. Available at https://cs231n.github.io/convolutional-networks/ [Accessed February 1, 2019]

Lloyd H. Hughes, Michael Schmitt, Lichao Mou, Yuanyuan Wang, Xiao Xiang Zhu., 2018 Identifying corresponding patches in sar and optical images with a pseudo-siamese cnn. IEEE Geoscience and Remote Sensing Letters 15.5 pp 784-788.

Muckenhuber, S., Korosov, A. A., Sandven, S 2016 Open-source feature-tracking algorithm for sea ice drift retrieval from Sentinel-1 SAR imagery. The Cryosphere, Volume 10, pp 913-925

QGIS Development Team 2009, QGIS Geographic Information System, Open Source Geospatial Foundation. Available at http://qgis.org. [Accessed February 1, 2019]

Simonyan, K, Zisserman., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Varior R. R., B. Shuai, J. Lu, D. Xu, and G. Wang, 2016. A siamese long short-term memory architecture for human re-identification. 14th European Conference Computer Vision, Amsterdam, The Netherlands, pp. 135–153.