Neural networks and satellite images-based shrub tundra landscape study: phenomena with fuzzy geometric and categorical boundaries

We apply Convolutional Neural Networks (CNNs) on submeter resolution satellite imagery acquired 10–15 years apart to map age stages of tall shrub growth and its changes over time for several sites in Low Arctic tundra landscapes.

Anna Derkacheva
Intl. Lab. of Landscape Ecology
HSE University
Moscow, Russia
aderkacheva@hse.ru | der_a@mail.ru

Gerald V. Frost
Dept. of Env.Sciences
University of Virginia
Charlottesville, USA

Ksenia Ermokhina
Severtsov Inst. of Ecology and Evolution
Russian Academy of Sciences
Moscow, Russia

Howie Epstein
Dept. of Env. Sciences
University of Virginia
Charlottesville, USA

Study phenomena

The expansion of shrubs in tundra ecosystems — "*shrubification*" — is among the most conspicuous and pervasive changes being observed in the terrestrial Arctic and is expected to continue in the future. Shrubification is thought to be a key driver of circumpolar "greening" trends observed by satellites and has far-reaching implications for the biophysical properties and function of Arctic ecosystems, including surface energy balance, regional climate, permafrost, and carbon cycling.

For the further comprehensive ecological study of shrubification, we aim to map extensively the spread of shrubs and thier spatial propagation over time.

We target four classes of alder shrub cover to map with CNNs: background tundra without alder, colonization stage with young alder individuals, middle stage with cover intensification and/or shrubs maturation, and latter stage with dense shrub cover.

... and its specificity

The typical objects of interest have obvious contours and categories of belonging in natural images for which the CNNs were developed (see such contest-used datasets as MNIST, ImageNet, OpenImages), and in a number of existing studies using satellite imagery. However, many natural objects, phenomena, and their properties have fuzzy gradual boundaries. Shrubification is such a "vague" target: it does not have clear division with regard to categorical classes of shrub growth stages, as well as clear geometric limit for each stage. Those boundaries can be even more blurred by the resolution of a satellite image.

Study area and Data used

For the first attempt, we selected three tundra-dominated locations across northwest Siberia in Russia: nearby Kharp, Obskaya, and Dudinka.

The dominant tall shrub at these sites is Siberian alder (*Alnus alnobetula*), a distinctive species that achieves heights up to 3 m and has a round-shape crown. It starts to grow with sparse individuals and can later form densely covered stands.

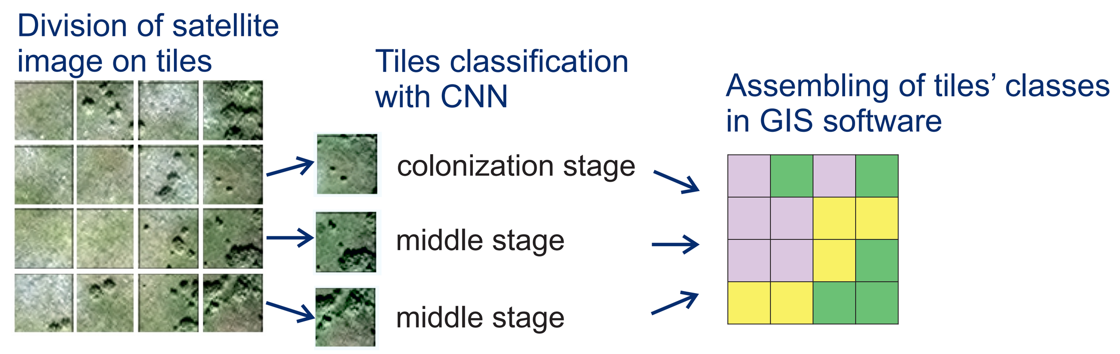
In some places it is mixed with larch trees, which have a particular triangular-shaped crown.

We use sub-meter resolution satellite imagery with four spectral bands (Red, Green, Blue, NIR) taken under sunny summer conditions. For each site, the images covering about 60-80 km2 were acquired for two time periods (see Table).

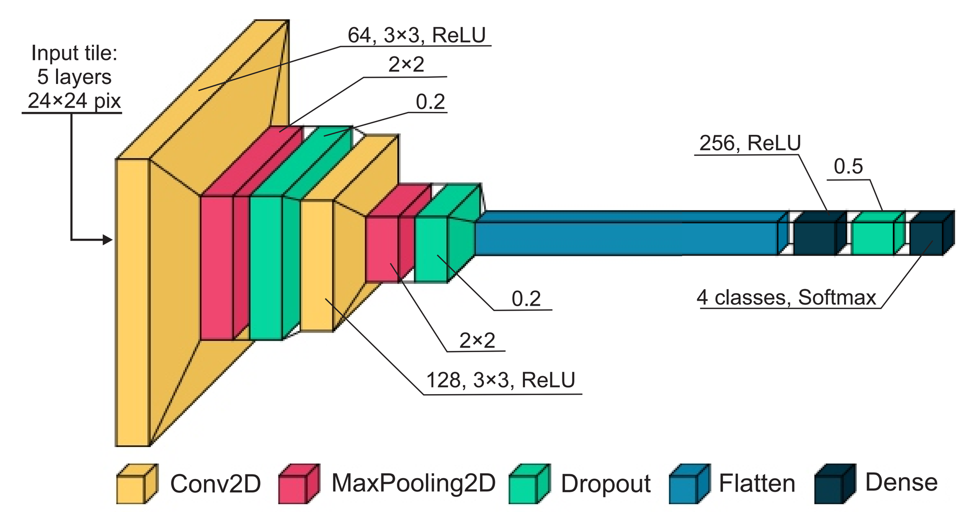
Methodology

We adopted an approach of "tile classification". The core of the processing is the *classification*, the easiest task for neural networks. Classification is applied independently to a number of small subimages (tiles) composing a regular grid, and the final results are combined all together to construct a full coverage map of the alder growth stage.

In our specific cases, the tiles are squares of 12 m on a side on the ground, which is about 20-40 pixels on a side, depending on the image resolution. This size is presumed to be a compromise between the detail of the final map and the loss of spatial context within a tile.



All steps are coded within the open-source Python libraries Keras/TensorFlow for working with CNN and Scikit-Learn for manipulating with annotated dataset, and run on GoogleColab. This is except for labeling the training tiles, which is easy and fast to do with GIS. The input tiling grid and output classification maps are ESRI shape files.



Our CNN has only two blocks with the convolutional layers, this is a limitation of the small size of the input tiles. The training process is controlled by the Categorical Cross-Entropy loss function and Adaptive Moment Estimation optimization method (Keras/ADAM). The duration of the training is limited to 100 epochs, and it can be interrupted by an early-stop option if the losses no longer decrease for 10 consecutive epochs. In addition to the training and validation datasets, the class weights are provided at the input for handling an issue of unbalanced classes.

Since the actual size of the input tiles and the image properties (resolution, sun conditions, etc.) vary from case to case, at the current stage of the study we decided to process each image independently, namely creating an annotated dataset, training CNN, and evaluating it.

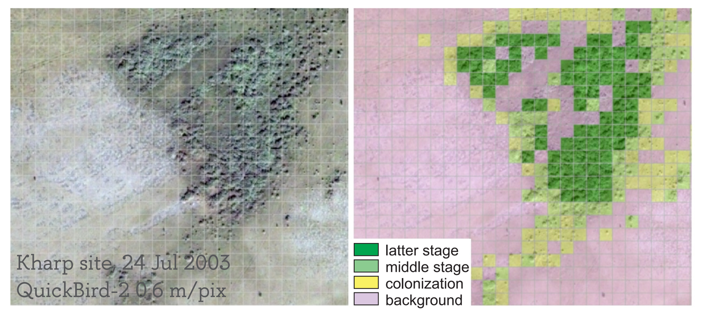
About 7'000-10'000 training tiles were created per image, with relatively balanced representation of three alder stages and more numerous background tiles (see Table). The number of samples is decided by an expert depending on the landscape complicity and heterogeneity, but should be above 1'000 for each class. Note that with GIS, it takes only a few hours to annotate such a training data set.

Results

We achieved the F-score to be above 0.8 for all 6 images and above 0.9 in the best cases.

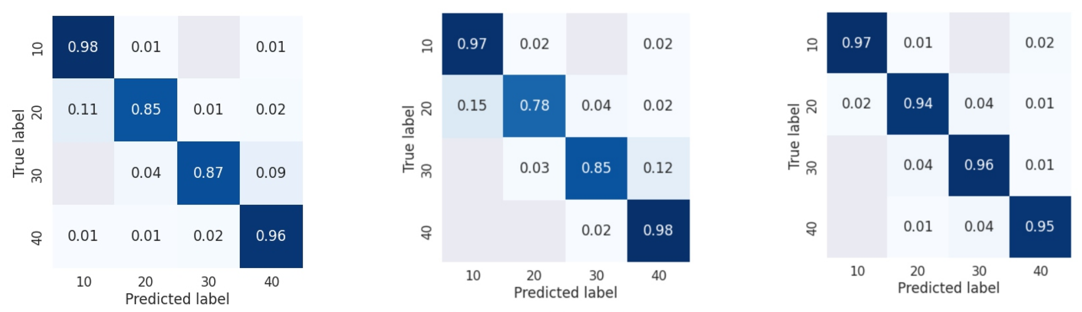
					Annotated dataset				CNN learning validation			CNN evaluation		
Site	Year	Image resolution, m/pix	Tile size,m (pix)	Entire image, tiles	Latter state	Middle state	Coloni-sation	Back-ground	R	P	F	R	P	F
Kharp	2003	0.6	12 (20)	441 557	1509	1046	1091	6166	0.897	0.901	0.899	0.906	0.910	**0.908**
	2015	0.3	12 (39)	337 939	1153	1451	1068	4765	0.825	0.840	0.833	0.831	0.848	**0.839**
Obskaya	2004	0.8	12 (16)	476 159	1362	1407	1040	3959	0.876	0.882	0.879	0.876	0.883	**0.879**
	2019	0.3	13 (42)	228 055	1361	1487	1349	3713	0.824	0.829	0.827	0.835	0.841	**0.838**
Dudinka	2009	0.5	12 (24)	523 443	1295	1776	1565	5918	0.893	0.896	0.894	0.881	0.883	**0.881**
	2019	0.5	12 (24)	411 863	1303	1573	1348	4209	0.922	0.925	0.924	0.922	0.923	**0.923**

R – Recall, P – Precision, F = 2PR/(P+R)



We note that the errors are not equally distributed across the shrubification classes, as well as spatially across the landscape. Regarding the latter, areas with many small ponds that appear as tiny round dark spots, and mixed stands of larch trees and alder shrubs are usually more affected by misclassification compared to dry areas or dense stands of alder.

Investigating the error dependency on classes, we expectedly found that the colonization class was the most error-prone. It was commonly misclassified as both false positives and false negatives with the background tundra class, which includes various forms of the landscape. Several problems affecting imagery reflectance, such as glare, also influence the colonization class more than others. At the same time, the middle and latter stages were much more readily distinguished from the background, and the misclassifications were relatively rare.



Obskaya site, 29 Jul 2004
GeoEye-1 0.8 m/pix

Kharp site, 24 Jul 2003
QuickBird-2 0.6 m/pix

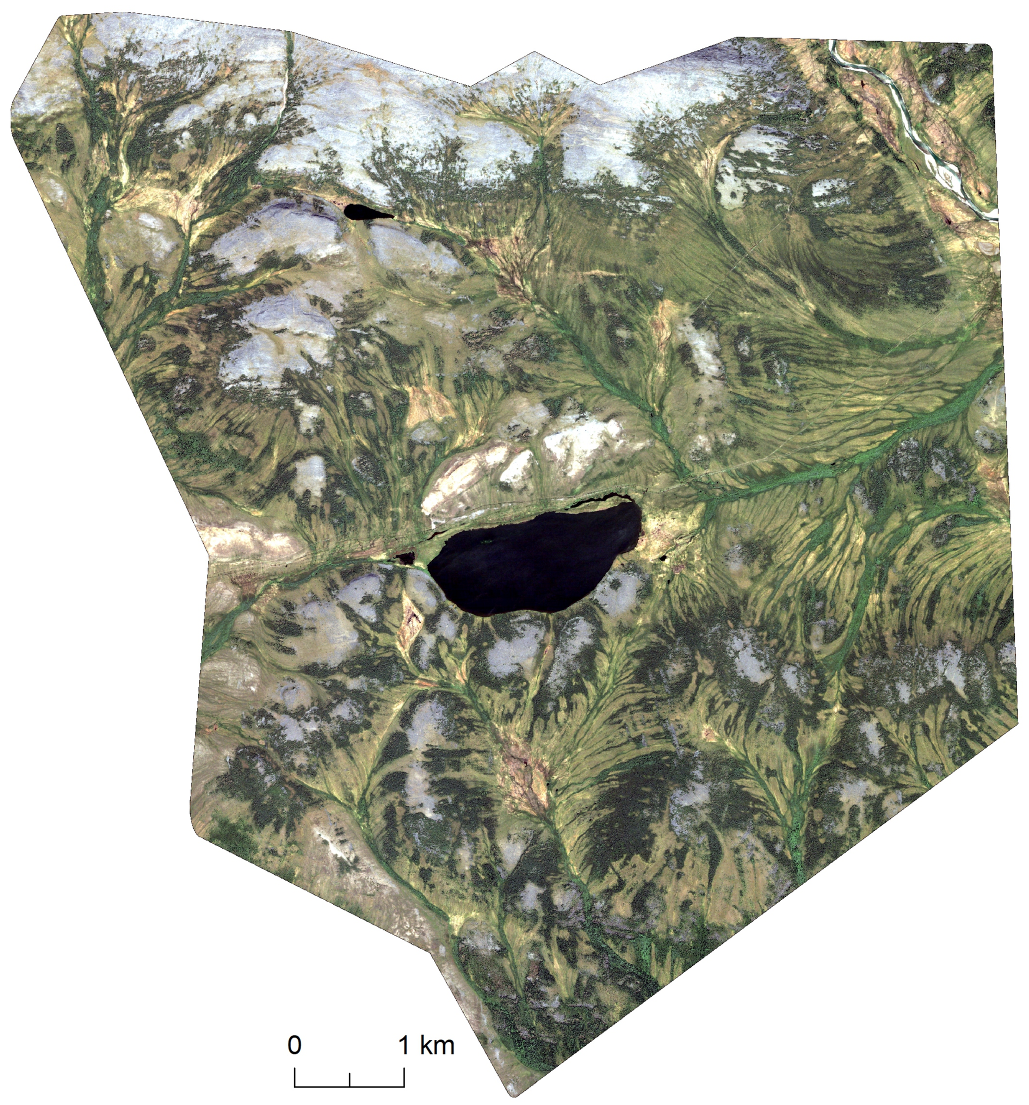
Dudinka site, 9 Jul 2009
GeoEye-1 0.5 m/pix

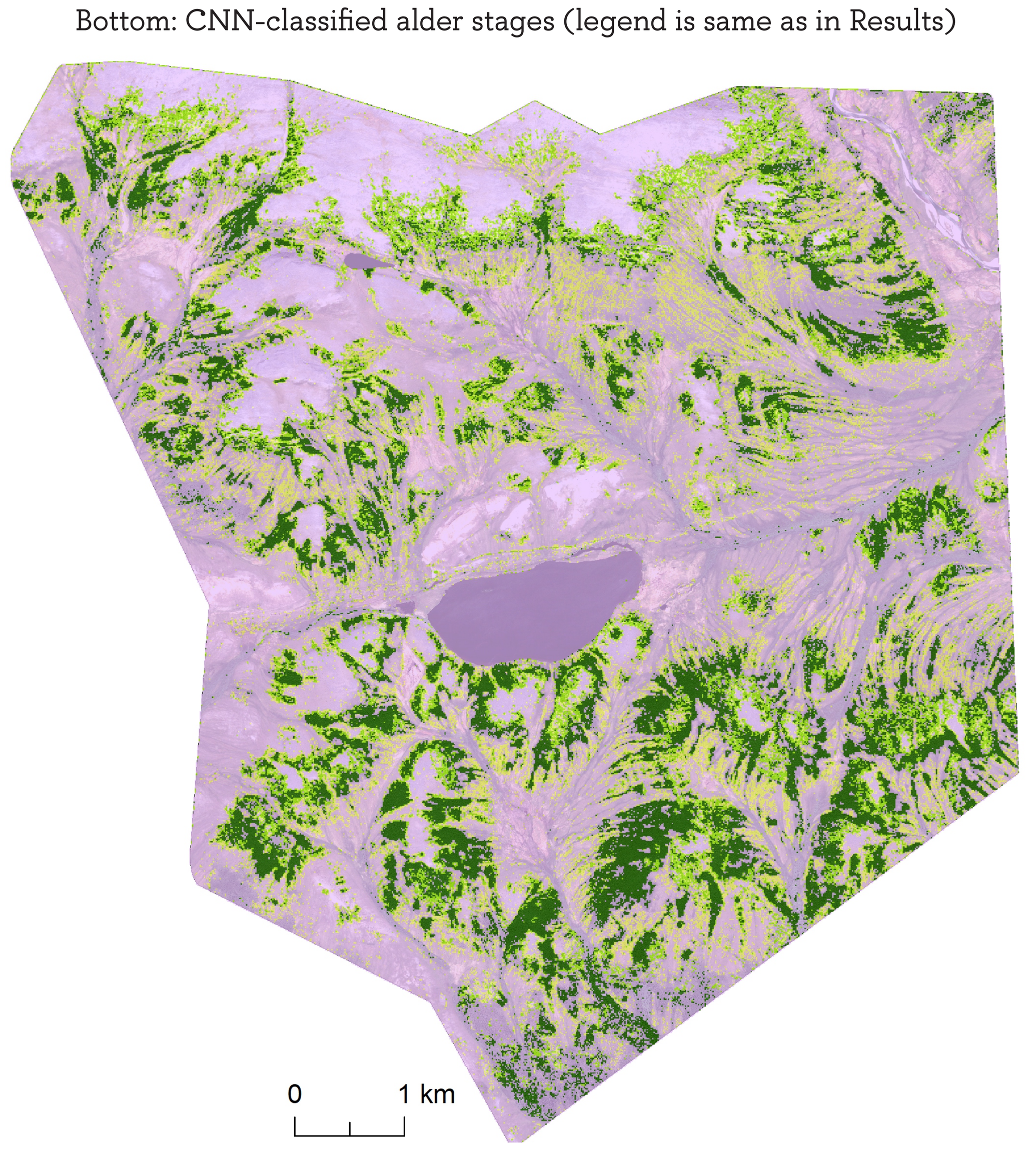
Label codes: 10 - latter stage with dense shrub cover, 20 - middle stage with cover intensification and/or shrubs maturation, 30 - colonization stage with young alder individuals, 40 - background tundra without alder

To investigate deeper the classes particularity, an independent experiment was conducted for several images. The networks with the similar architecture were trained and evaluated using an isolated alder class against the background (e.g. only colonization and background tiles). The evaluation F-scores obtained in such cases were typically as follow: 0.90-0.95 for the latter stage, 0.80-0.85 for the middle stage, and 0.60-0.70 for the colonization stage. The second test consisted in training of a CNN to predict only two classes: presence of alder (all stages together) and background. The evaluation F-scores for this test were between 0.86 and 0.96.

What is next

The next stage of our work is the analysis of the derived data. Along with assessments of the manifestation and intensity of shrubification in selected locations, we plan to compare the obtained maps with a number of local environmental parameters (e.g. rocks state, slope aspect, land form, etc.) in order to investigate conditions favourable for the spread of shrubs. The second point to be addressed, is the increase in the number of processed images and covered locations. This will not only bring the additional thematic data, but will also allow for creation of a robust processing chain, capable to deal with diverse images and landscapes without CNN training step.





Kharp site
Top: satellite imagery QuickBird-2, 0.6 m/pix, 24 Jul 2003
Bottom: CNN-classified alder stages (legend is same as in Results)

Obskaya site (fragment)
Left: satellite imagery WorldView-3, 0.3 m/pix, 25 Jul 2019
Center: CNN-classified alder stages (legend is same as in Results)
Right: comparison of 2004 and 2019 (satellite imagery on background)