2023 International Conference on Machine Intelligence for GeoAnalytics and Remote Sensing (MIGARS) | 979-8-3503-4542-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/MIGARS57353.2023.10064585

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

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

Gerald V. Frost
Alaska Biological Research, Inc.
Fairbanks, USA
jfrost@abrinc.com

Ksenia Ermokhina
Severtsov Inst. of Ecology and Evolution
Russian Academy of Sciences
Moscow, Russia
diankina@gmail.com

Howie Epstein
Dept. of Environmental Sciences
University of Virginia
Charlottesville, USA
hee2b@virginia.edu

***Abstract*—Many studies using convolutional neural networks, including in the field of satellite images, are aimed at recognizing clearly defined objects, such as cars or individual trees. Here we present the first results on mapping the *growth stage of tundra shrubs*, which is a "fuzzy" target for a network: there are no obvious geometric boundaries of the growth stages, as well as a clear division between the classes representing the growth stages. For six high-resolution satellite images of three Low Arctic landscapes, we achieved an F-score of 0.8-0.9, what is very promising for further ecological research. We also found that different landscapes and stages of shrub growth are not equally error-prone.**

***Index Terms*—Earth remote sensing, Convolutional Neural Network, tundra landscape, ecology**

I. Introduction

A number of natural phenomena, objects, and landscape characteristics cannot be detected with Earth-observing satellite imagery using traditional spectral and segmentation processing methods. This forces those who are studying such objects and phenomena to conduct visual interpretations of images, limiting the spatial and temporal extents of their study. Given the current exponential growth in the quantity of available imagery, this leads to a huge underutilization of information for environmental monitoring, modeling, and forecasting. This is especially true for historical imagery sources. The development of machine learning algorithms and artificial intelligence has enabled advanced forms of imagery analysis with unprecedented spatial resolution and scalability. Image processing methods based on neural networks are able to cope with both issues: 1) complex and/or abstract objects that are unrecognizable with classic image-processing methods, and 2) large volumes of satellite images that need to be processed rapidly.

In this work we apply Convolutional Neural Networks (CNNs, [1], [2]), a deep machine learning technique, 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.

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 [3]. Shrubification is thought to be a key driver of circumpolar "greening" trends observed by satellites [4] and has far-reaching implications for the biophysical properties and function of Arctic ecosystems, including surface energy balance, regional climate, permafrost, and carbon cycling. Increasing Arctic summer temperatures due to global warming are thought to be the central driver of shrubification.

CNNs hold promise for filling knowledge gaps and overcoming strong geographic biases in our current understanding of the scale and pace of shrubification and other forms of Arctic environmental change.

It is important to note that the typical objects of interest have obvious contours and categories of belonging in natural images for which the CNNs were developed (see first-referring paper [1] and such contest-used datasets as MNIST, ImageNet, OpenImages), and in a number of existing studies using satellite imagery (see reviews [5], [6]). 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.

The article was prepared within the framework of the Basic Research Program at HSE University, Moscow, RF. The satellite images were obtained thanks to the Department of Environmental Sciences at the University of Virginia, Charlottesville, USA

Authorized licensed use limited to: Higher School of Economics. Downloaded on March 22,2023 at 07:27:01 UTC from IEEE Xplore. Restrictions apply.

II. Study area and data

For the first attempt, we selected three tundra-dominated locations across northwest Siberia in Russia: nearby Kharp, Obskaya, and Dudinka (Fig.1). Those sites have previously been investigated with regard to warming-induced shrubification, using field methods and visual analysis of repeating high-resolution satellite images [7], [8]. 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, which has a particular triangular-shaped crown.

Fig. 1. Location of the selected sites.

We acquired satellite imagery for the three sites for two periods (see Table I) under sunny summer conditions. GeoEye-1, QuickBird-2, IKONOS and WorldView-2 satellites cover about 60-80 km^2 per site, with four bands (Red, Green, Blue, NIR) at sub-meter resolution after pan-sharping done. On those images, individual alders look like circles with different sizes and colors related to illuminated and shaded sides (Fig.3, left), and typically cast distinct round shadows on the underlying tundra vegetation. The dense alder stands exhibit a spongy structured pattern.

III. Methods

A. Processing pipeline

For our task, we applied the CNNs, a class of artificial neural networks designed and commonly used for image pattern recognition [1]. The tasks currently approached with CNNs are commonly divided into four groups: *classification* of the entire image with one resulting label, *object detection* that marks the object location, *image segmentation* with a final graphical result similar to a traditional pixel-based classification, and *instance segmentation* that identifies the exact shape and class of each recognised object [9].

To deal with the issue of the fuzzy geometric boundaries of shrubification, we developed 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 (Fig.3). Although we came up with this idea on our own, it has already been invented and applied in satellite image processing tasks. For instance, Kattenborn et al. [10] used exactly the same approach to map plant species with a target characteristic of cover fraction in a tile. Marochov et al. [11] integrated this approach in a more sophisticated processing chain to obtain a land cover map of the Greenlandic glaciers margin. Open-access datasets such as EuroSAT [12] and BigEarthNet [13], [14] for Land Use/Land Cover mapping are also organized as labeled tiles.

Our processing is enveloped by a start-to-end pipeline optimized for earth scientists who commonly use Geographic Information Systems (GIS) for mapping. 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.

First of all, a shape file, containing a grid that covers an entire study area, is generated as an image tile-cropping simulation with the required tile size. A set of tiles, arbitrarily chosen by a user, is labeled according to shrub occurrence and successional stage in the grids's attribute table with any shape-file friendly GIS package. This annotated dataset is further randomly divided into the training, validation and evaluation subsets with a ratio of 60:20:20. Then, a CNN based on linear ResNet architecture is trained on a subset of those labeled tiles. At this one and classification steps, a Normalized Difference Vegetation Index (NDVI) image is computed from the original Red and NIR bands as an additional input layer to improve the classification quality. Finally, the trained CNN is applied to classify the tiles across the entire study area grid, with output class and probability stored in the grid's attribute table (Fig.3).

At all stages, the shape file grid is used for "on-the-fly" tile manipulations including NDVI-image generation, instead of storing them to physical media. Such a design saves memory and accelerates the processing by deprecating repetitive read-write operations, as the average size of our input images is hundreds of thousands of tiles (Table I). Usage of the same shape file grid for labelling, learning and processing allows for easy organization of quality assessment.

To estimate the classification quality, considering the imbalanced multiclass nature of our problem, we use the confusion matrix as well as the metrics of Precision P, Recall R, and F-score F=2PR/(P+R). The metrics vary between 0 and 1 with the latter being the best result.

B. Technical realization

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. The architecture of our CNN is presented in Fig.2. It has only two blocks with

Authorized licensed use limited to: Higher School of Economics. Downloaded on March 22,2023 at 07:27:01 UTC from IEEE Xplore. Restrictions apply.

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.

We target four classes of alder shrub cover: *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 (Fig.3, right). About 7'000-10'000 training tiles were created per image, with relatively balanced representation of three alder stages and more numerous background tiles (Table I). 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.

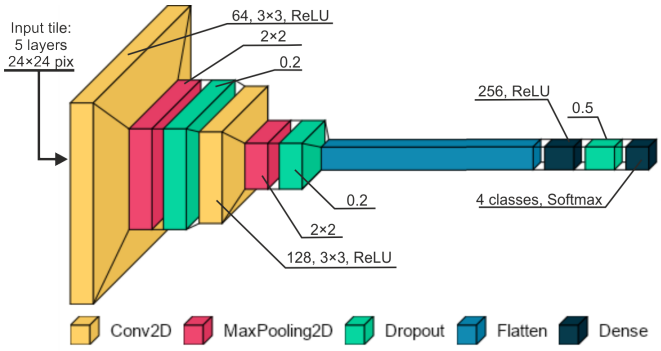

Fig. 2. Architecture of the used CNN model (actual size of the input tiles vary from image to image).

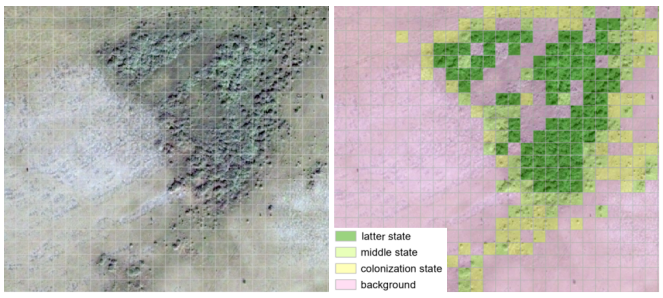

Fig. 3. Left: source QuickBird satellite image for Kharp site (0.6 m/pix, 2003 yr); gray lines are the 12 m tiling grid. Right: classification for the same area done by CNN.

IV. Preliminary Results

We achieved the F-score to be above 0.8 for all 6 images and above 0.9 in the best cases. The exact scores of quality assessment (validation during the CNN training and evaluation of a trained CNN) are given in Table I.

Regarding the misclassifications done by CNN, outputs suffer from both types of errors: *false positive* meaning no alder exists but is detected by CNN, and *false negative* meaning that an existing alder is not detected. However, 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 (Fig.4). 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. 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.

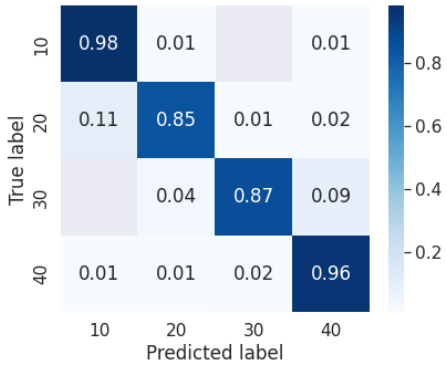

Fig. 4. Classes confusion matrix (fraction) for Obskaya site (source image - GeoEye-1, 0.8 m/pix, 2004 yr). Classes are coded as follow: 10 - latter stage, 20 - middle stage, 30 - colonization stage, 40 - background.

Note that while the misclassification of alder against the background is an error purely induced by CNN due to network quality or due to imagery limitations (mainly, ground resolution), the cross-attribution of alder stages comes partly

Authorized licensed use limited to: Higher School of Economics. Downloaded on March 22,2023 at 07:27:01 UTC from IEEE Xplore. Restrictions apply.

TABLE I
PARAMETERS OF THE PROCESSED IMAGES, SIZE OF ANNOTATED DATASETS, AND CNN TRAINING SCORES

					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**

from the expert-prepared training data. This is due to the fuzzy classes and geometric boundaries for each class. Therefore, the training dataset itself is subject to the effects of expert classification "drift", meaning that similar tiles could be classified differently by different people, or even by the same person depending on the surrounding landscape or state of mind during the process. We conducted a brief experiment with our team of co-authors and found the discrepancies up to 10% in the most complex landscapes. The majority of disagreement occurred between colonization and middle stages.

V. CONCLUSION

CNNs are a promising instrument for widespread high-resolution mapping of landscape characteristics in Earth science. Here we demonstrated that CNN can successfully be used to study Arctic shrubification, in which the landscape features of interest have fuzzy geometric and categorical boundaries. The designed processing chain includes the splitting of an image into numerous tiles, the independent classification of each tile by a CNN according to the shrubs growth stages, spatial assembling of the results to get a final map. This processing chain consistently showed F-scores (a mix of precision and recall) above 0.8 at six very-high-resolution images across three different Arctic locations.

Our study also revealed that the positive and negative misclassifications are distributed unequally among the landscape types and shrubification classes. The former are mainly errors induced by CNN that are related to complex landscape features. The latter is induced not only by CNN, but also by the subjectivity of experts preparing the training and validation datasets, as a consequence of fuzzy categorical boundaries.

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 favorable 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.

REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2012.

[2] A. Ghosh, A. Sufian, F. Sultana, A. Chakrabarti, and D. De, "Fundamental Concepts of Convolutional Neural Network," pp. 519–567, 2020.

[3] I. H. Myers-Smith, B. C. Forbes, M. Wilmking, M. Hallinger, T. Lantz, D. Blok, K. D. Tape, M. MacIas-Fauria, U. Sass-Klaassen, E. Lévesque, S. Boudreau, P. Ropars, L. Hermanutz, A. Trant, L. S. Collier, S. Weijers, J. Rozema, S. A. Rayback, N. M. Schmidt, G. Schaepman-Strub, S. Wipf, C. Rixen, C. B. Ménard, S. Venn, S. Goetz, L. Andreu-Hayles, S. Elmendorf, V. Ravolainen, J. Welker, P. Grogan, H. E. Epstein, and D. S. Hik, "Shrub expansion in tundra ecosystems: Dynamics, impacts and research priorities," *Environmental Research Letters*, vol. 6, no. 4, 2011.

[4] G. V. Frost, U. S. Bhatt, H. E. Epstein, I. Myers-Smith, G. K. Phoenix, L. T. Berner, J. W. Bjerke, B. C. Forbes, S. J. Goetz, J. Kerby, M. J. Macander, T. Park, M. K. Raynolds, H. Tømmervik, and D. Walker, "Tundra Greenness," 2020.

[5] T. Hoeser, F. Bachofer, and C. Kuenzer, "Object detection and image segmentation with deep learning on earth observation data: A review-part II: Applications," *Remote Sensing*, vol. 12, no. 18, 2020.

[6] P. G. Brodrick, A. B. Davies, and G. P. Asner, "Uncovering Ecological Patterns with Convolutional Neural Networks," *Trends in Ecology and Evolution*, vol. 34, no. 8, pp. 734–745, 2019.

[7] G. V. Frost, H. E. Epstein, D. A. Walker, G. Matyshak, and K. Ermokhina, "Patterned-ground facilitates shrub expansion in Low Arctic tundra," *Environmental Research Letters*, vol. 8, no. 1, 2013.

[8] G. V. Frost and H. E. Epstein, "Tall shrub and tree expansion in Siberian tundra ecotones since the 1960s," *Global Change Biology*, vol. 20, no. 4, pp. 1264–1277, 2014.

[9] T. Hoeser and C. Kuenzer, "Object detection and image segmentation with deep learning on Earth observation data: A review-part I: Evolution and recent trends," *Remote Sensing*, vol. 12, no. 10, 2020.

[10] T. Kattenborn, J. Eichel, S. Wiser, L. Burrows, F. E. Fassnacht, and S. Schmidtlein, "Convolutional Neural Networks accurately predict cover fractions of plant species and communities in Unmanned Aerial Vehicle imagery," *Remote Sensing in Ecology and Conservation*, vol. 6, no. 4, pp. 472–486, 2020.

[11] M. Marochov, C. R. Stokes, and P. E. Carbonneau, "Image classification of marine-terminating outlet glaciers in Greenland using deep learning methods," *The Cryosphere*, vol. 15, pp. 5041–5059, nov 2021.

[12] P. Helber, B. Bischke, A. Dengel, and D. Borth, "EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, pp. 2217–2226, jul 2019.

[13] G. Sumbul, M. Charfuelan, B. Demir, and V. Markl, "BigEarthNet: A Large-Scale Benchmark Archive for Remote Sensing Image Understanding," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 5901–5904, IEEE, jul 2019.

[14] G. Sumbul, A. de Wall, T. Kreuziger, F. Marcelino, H. Costa, P. Benevides, M. Caetano, B. Demir, and V. Markl, "BigEarthNet-MM: A Large-Scale, Multimodal, Multilabel Benchmark Archive for Remote Sensing Image Classification and Retrieval," *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, pp. 174–180, sep 2021.

Authorized licensed use limited to: Higher School of Economics. Downloaded on March 22,2023 at 07:27:01 UTC from IEEE Xplore. Restrictions apply.