Using deep learning algorithms for texture segmentation of ultra-high resolution satellite images

Dmitry Rusin[1]*, *Anna* Alehina[1], *Anastasia* Safonova[1], and *Egor* Dmitriev[2]

[1] Deep Learning Laboratory, Siberian Federal University, 660041 Krasnoyarsk, Russia
[2] Marchuk Institute of Numerical Mathematics of the Russian Academy of Sciences, 119991 Moscow, Russia

Abstract. This paper presents the results of textural segmentation of satellite images with spatial resolution <1 m using U-Net convolutional neural networks. To conduct numerical experiments, a panchromatic image of the WorldView-2 test site on the territory of the Bronnitsky Forestry (Moscow region) used. The possibilities of automating the selection of neural network parameters based on genetic algorithms investigated. The proposed method makes it possible to effectively segment the main types of natural and man-made objects, as well as to distinguish structural classes of woodlands.

1 Method

In this paper, the task was set to create a machine-learning model based on convolutional neural networks. This model should automatically segment natural and technogenic objects on panchromatic satellite images of ultra-high spatial resolution by textural features. The *Xception* model used in the paper [1] with the ability to change a number of parameters to improve accuracy. Architecture *Xception* is a convolutional neural network that is capable of working with a small amount of training data for segmentation tasks. The generalized *U-Net* architecture [2] shown in Figure 1.

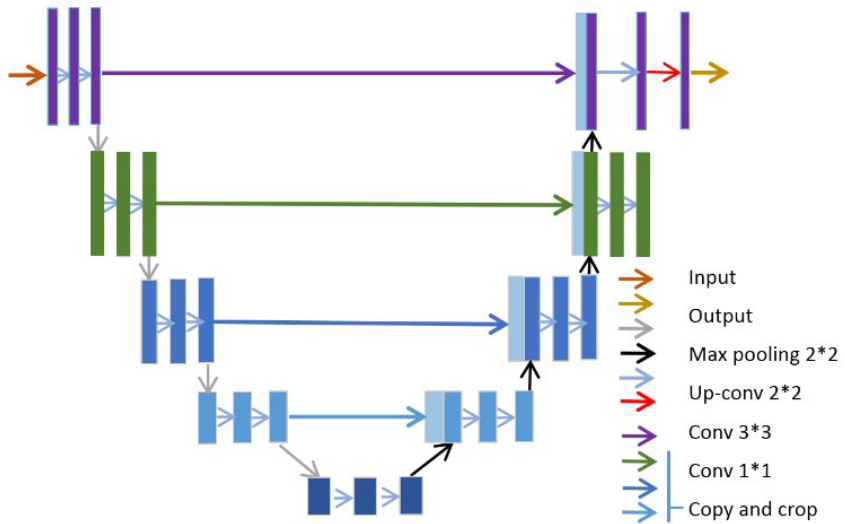* Corresponding author: drusin@sfu-kras.ru

Fig. 1. Generalized U-Net architecture (figure from https://arxiv.org/abs/1505.04597).

Automated parameter setting performed using a genetic algorithm (GA). The algorithm based on the output quality assessment of the U-Net model. Calculations performed to assess the effectiveness of automated parameter tuning. The applied model trained in two ways, with the use of GA and without it. To compare the quality of the models, it decided to use cross-qualification [3] for evaluation on independent data. Cross-qualification applied only on the standard U-Net model without GA. The GA parameters presented in Table 1.

Table 1. GA parameters.

Name	Meaning
Mutation coefficient	0.1
Population	20
Number of iterations	50
Crossing	3

2 Data

The data set consists of satellite images of the Bronnitsky forestry of the Moscow region, the total image has a size of 10140 x 10120. First, this image processed: it divided into 64 parts for subsequent data markup.
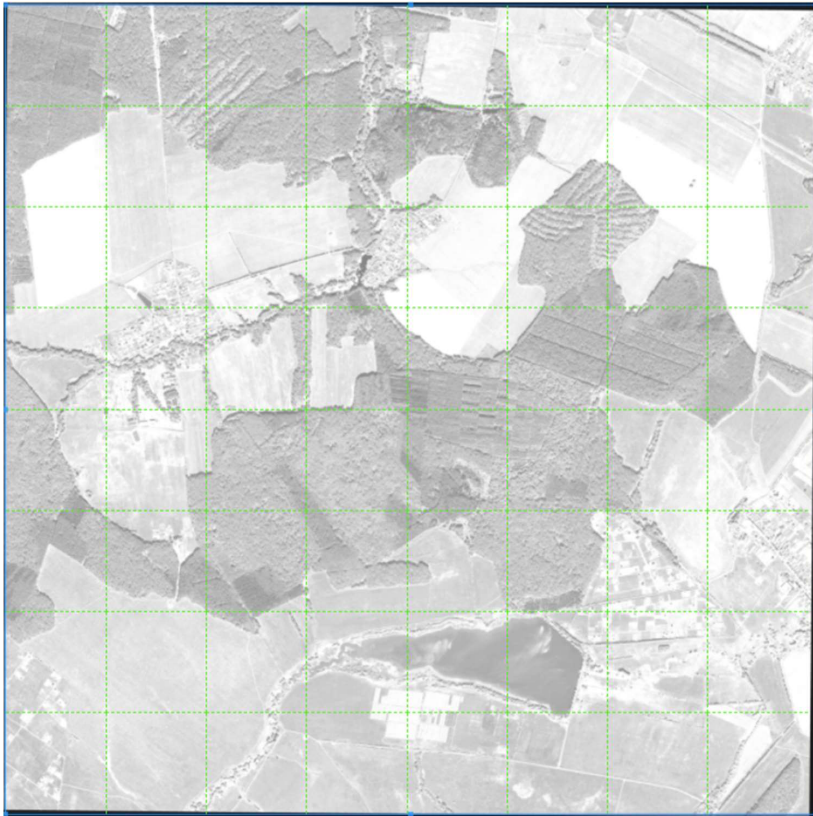
Fig. 2. Splitting the source image into frames.

Splitting the snapshot allows you to reduce the load during training, as well as mark up the data in more detail. The final size of each frame is 1267x1265.

In this task, six classes were allocated: mixed forest, field, city, forest belt, ordered forest, water.

To assess the quality of the model, the mAP (4) metric used, based on the error matrix F-measure (3), it also represents a balance between two metrics Precision (1) and Recall (2). Categorical cross-entropy (5) used as the loss function of the model.

$$Precision = \frac{TP}{TP+FP}$$

(1)

$$Recall = \frac{TP}{TP+FN}$$

(2)

$$F\_score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{3}$$

$$mAP = \sum_{i=1}^{N} AP_i = \frac{1}{N} \sum_{Recall} \mathrm{Pr}ecision(\mathrm{Re}call_i) \tag{4}$$

$$Loss = - \sum_{i=1}^{Output\,size} y_i * \log y_i \tag{5}$$

where, y_i is the scalar value at the output of the model, output size is the number of scalar values in the output of the model.

An example of the markup of one of the frames shown in Figure 3.

Fig. 3. Example of a marked-up snapshot.

3 Results

The neural network architecture (NN) trained using a free Google Colab cloud-computing machine [4]. The main results of the experiment on test data presented in Table 2.

Table 2. Results of the work

Name of the metric	Unet	Unet+GA	Cross-qualification (folds = 5)
F-measure	75.74	77.86	82.37
Loss	22.12	16.75	13.00
mAP	71.54	73.56	76.49
Time spent on training (1 iteration, min)	*25*	55	130

The results of the work of U-Net and U-Net+GA display an average of 15 launches each. As can be seen from the table, the U-Net model copes with the processing of satellite images in the task of segmentation. For higher results, a more precise network setup needed, since even changing several parameters allowed to increase the accuracy of the model. In addition, the model is sensitive to the amount of input data. This seen based on the results of cross-validation. In the future, it is planned to test these models on new sites.

References

1. F. Chollet, Xception: Deep Learning with Depthwise Separable Convolutions, arXiv:1610.02357v3 [cs.CV], 2017
2. O. Ronneberger, U-Net: Convolutional Networks for Biomedical Image Segmentation, arXiv:1505.04597, 2015
3. F. Mosteller and J.W. Tukey Data analysis, including statistics. In Handbook of Social Psychology. Addison-Wesley, Reading, MA, **46**, 1968.
4. Google Colab [Electronic resource] - Access mode: https://colab.research.google.com