**Wildfire Segmentation on Satellite Images using   
Deep Learning**

Vladimir Khryashchev, Roman Larionov

P.G. Demidov Yaroslavl State University

[v.khryashchev@uniyar.ac.ru](mailto:v.khryashchev@uniyar.ac.ru), [r.larionov@uniyar.ac.ru](mailto:r.larionov@uniyar.ac.ru)

ABSTRACT

In this article there is presented a convolutional neural network for automated wildfire detection on high-resolution aerial photos. The model was implemented by means of Keras library and launched on modern GPUs of high-performance supercomputer NVIDIA DGX-1. Various techniques of data augmentation are used to enlarge training and test sets generated by data windowing. Special metrics, such as Sorensen-Dice coefficient, precision, recall, -score and IoU value allows to measure the quality of developed model.

Keywords: computer vision, forest fire detection, image segmentation, satellite images

1. INTRODUCTION

A wildfire is a spontaneous, uncontrolled fire in areas of combustible vegetation occurring in rural areas. Earth is an flammable planet owing to seasonally dry climates, atmospheric oxygen, and widespread lightning and volcanic ignitions. However, last years the human factor has become the main reason of causing irreversible forest fires.

Strategies for wildfire prevention and suppression have varied over the years [1, 2]. Wildfire detection is carried out in five ways:

* observation from specially equipped fire observation towers and other structures;
* ground observation on foot or by vehicles;
* air surveillance using special instruments on aircrafts and helicopters;
* analysis of information from space;
* accounting of messages of locals.

The benefit of wildfires is the natural renewal of forests. However, prolonged fires change the composition of the air significantly. The main harm from forest fires is the depletion of flora and fauna, as well as the damage to natural resources. In addition, there is a reason for concern about the harm to human health, in particular to respiratory and circulatory systems. In 2010 American Heart Association published a scientific statement stating that there is a link between air pollution from tiny particles that appear in the air as a result of wildfires and cardiovascular diseases.

On Earth, more than 340 million hectares of forests are annually damaged by fire. The largest areas of burning are in Australia and African countries [3]. According to statistics, Russia takes 8th place among all countries of the world which have the most total area of forests destroyed by fire. Thus, there is urgently needed to monitor forests and detect wildfires in the beginning of their spreading. The size of hotbed of fires makes it possible to detect them from space by means of artificial intelligence methods applied to satellite images.

Remote sensing cannot completely replace methods of ground-based data collection, but it can help in monitoring of huge areas and hard-to-reach regions. Despite the rapid development of computer vision algorithms for detecting objects, automated image segmentation has not reached to same quality as manual marking of high-resolution aerial photos. Although a human is able to solve a segmentation problem better than a computer, it takes too much time which means that it is impossible to obtain results in real time. Therefore, the task of satellite image segmentation using computer vision algorithms is particularly relevant.

Image segmentation is a challenging task. Nowadays, various machine learning algorithms for detecting objects on satellite images exist. The main approach of these models is marking image pixels to corresponding classes of objects. The greatest effeciency of solving segmentation problem is achieved by using convolutional neural networks (CNNs) [4].

This paper presents a convolutional neural network that can be used for forest segmentation. The training process, testing and special metrics for assessing the quality of neural network work are described. Our work continues research, which was presented in [5, 6].

2. DATA PREPARATION

In our research, there were used the Resource and the Planet databases. The Resource dataset consists of 10-bit three-channel high-resolution aerial photos with a spatial resolution of 1 and 10 m/pixel. The Planet database contains 10-bit three-channel satellite images with a spatial resolution of 3 m/pixel. Some satellite images of these databases have noisy pixels, such as photographed clouds or glares from reflecting surfaces. To align aerial photos of both datasets satellite images with a low spatial resolution were reduced in size. Furthermore, all aerial photos were normalized: values of pixels were converted into the range [0, 255].

Satellite image segmentation concerns the usage of fragments of aerial photos, which are fed to the input of CNN, so before the training of developed CNN each satellite image and mask of datasets have been sliced on parts of 256 × 256 size with the step of 128 by data windowing. The intersection of patches allows to cope with problem of artifacts that occur at the junction of image fragments during the process of combining segmentation results. As a result, training and test sets of 1457 and 393 images and corresponding masks were formed. Examples of sliced images and masks are shown in Fig. 1. Every little part of sliced images corresponded to the needed small part of big generated mask. Information about prepared patches of different databases is presented in Tables 1, 2. Examples of sliced images and masks of the Resource and the Planet datasets are shown in Fig. 1, 2.

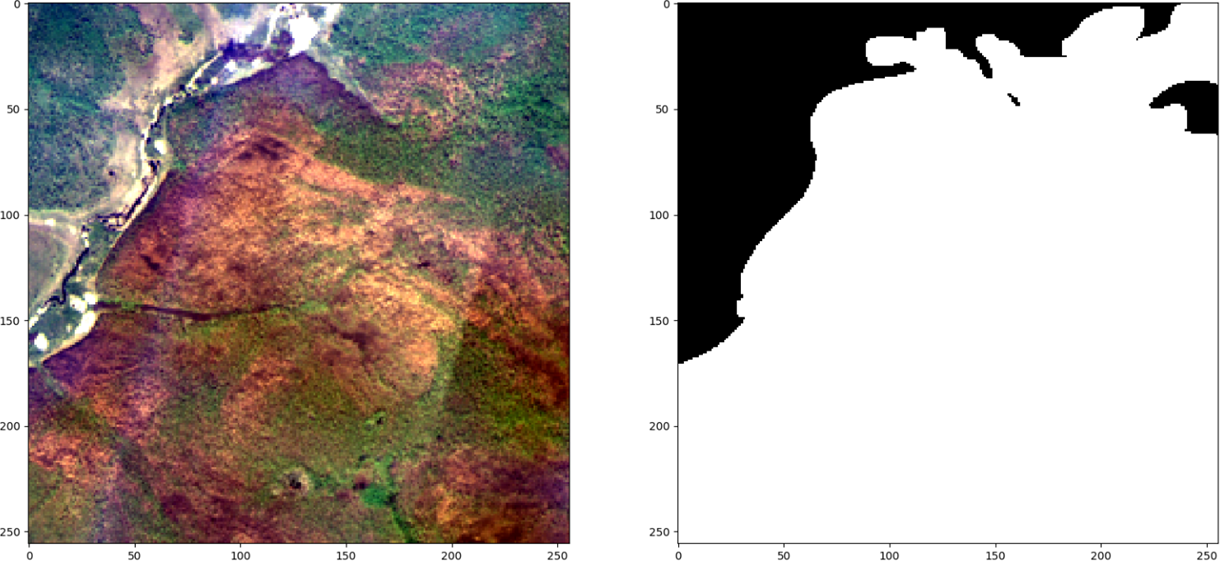
**Table 1. Prepared patches of the Resource database**

|  |  |  |
| --- | --- | --- |
|  | **Training set** | **Test set** |
| **Total** | 907 | 396 |
| **With objects** | 294 | 95 |
| **Without objects** | 613 | 301 |

**Table 2. Prepared patches of the Planet database**

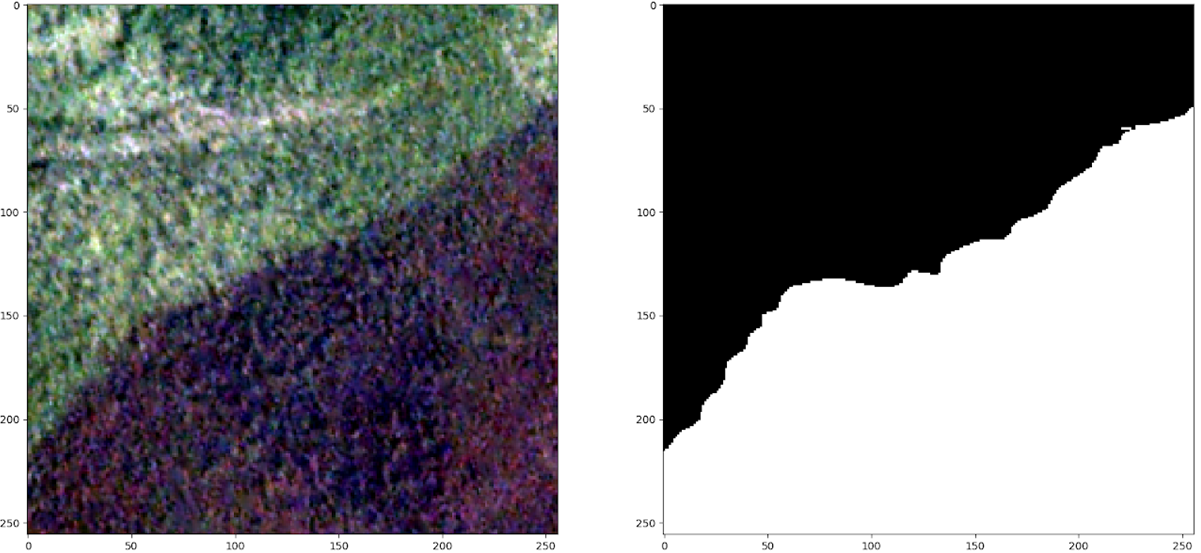
|  |  |  |
| --- | --- | --- |
|  | **Training set** | **Test set** |
| **Total** | 4591 | 3186 |
| **With objects** | 713 | 319 |
| **Without objects** | 3878 | 2867 |

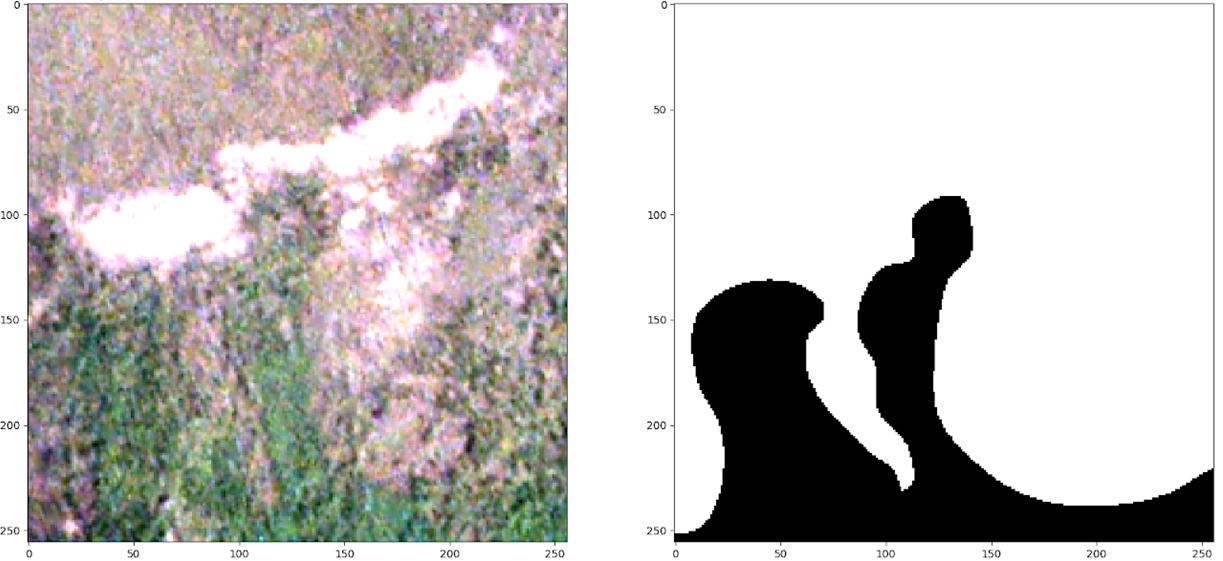




**a) b)**

**Figure 1. Examples of a sliced image and a corresponding mask of the Resource dataset:   
a) prepared patches, b) true masks**





**a) b)**

**Figure 2. Examples of a sliced image and a corresponding mask of the Planet dataset:   
a) prepared patches, b) true masks**

Generated patches and sliced masks obtained for the Resource database were not enough for training of developed CNN. Therefore, to enlarge training and test sets there were used the following techniques:

* rotations on 90˚, 180˚ and 270˚ and mirroring of patches. As a result, training and test sets were increased 8 times (flips);
* image shifts within 2% of image size, scaling on a coefficient from [1; 1,2] and rotations on small angles from   
  [-15˚, + 15˚] (SSR).
* applying random chromatic distortion in HSV color format. This type of data augmentation allows to increase the robustness of deep learning algorithm for noisy images, such as small clouds, glare from reflective surfaces (NSV\_dist).

3. CONVOLUTIONAL NEURAL NETWORK

The CNN, which was used for wildfire segmentation on satellite images in our research is based on very popular and widespread architecture called U-Net. This CNN was originally developed for segmentation of medical images. Its classical structure is described in [7]. For the first time U-Net showed its efficiency for the task of medical image segmentation. Later this model was applied to pixel-wise classification of satellite images. U-Net is an U-shaped CNN which consists of two parts: the encoder and the decoder. The encoder represents typical downsampling path of CNN, whereas the decoder represents upsampling path which is used for restoration of segmentation mask. The main peculiarity of this architecture is that it uses skip-connections to combine low-level and higher-level maps of features from the encoder and the decoder. The last layer of U-Net is convolutional layer with 1 × 1 filter, which allows to classify every pixel.

In our research there was developed UNet-like architecture – U-ResNet34 based on models from paper [8]. In this paper the authors hold the similar method to solve the problem of satellite images segmentation: they developed U-Net like architecture, which is using ResNet-34 weights in the encoder. This algorithm shows excellent results of detecting roads on satellite images of DeepGlobe database [9].

The approach based on the usage of CNNs requires considerable computational resources. Therefore, training and test stages were implemented on a large number of independent streams of GPU using the parallel computing technology NVIDIA CUDA. This cross-platform technology is supported by all modern NVIDIA graphics cards [10].

Keras library with Tensorflow framework as a backend was used for development of CNN. Keras is an open-source library written in Python. It is built on Tensorflow framework and contains various implementations of commonly used neural network building blocks, such as layers, activation functions and optimizers, and ready tools to preprocess images and text data. In other words, Keras offers a higher-level, more intuitive set of abstractions to develop deep learning models [11]. Moreover, this library allows to train and test developed models on GPU.

As an algorithm of numerical optimization Adam optimizer with learning rate of 1e-3 was chosen. This optimizer combines the best approaches from gradient descent and momentum optimizers and shows optimal and quick convergence for most tasks of machine learning [12].

4. NUMERICAL RESULTS

U-ResNet34 was launched on NVIDIA DGX-1 supercomputer, which was provided by Artificial Intelligence Center of P.G Demidov Yaroslavl State University.

As a rule, the quality of algorithms for image segmentation is evaluated by special metrics which compare predicted and true masks. In our research there was used Sorensen-Dice coefficient (), precision (), recall (), -score () and coefficient evaluated for true positive examples. The value of -score was used as the main metric of measuring the quality of developed CNN. -score combines recall and precision metrics, which show the ability of the algorithm to detect objects and the ability to distinguish classes from each other [13, 14]. -score is calculated by the following formula:

.

During the training, there was chosen a threshold value () with step 0,01 to form predicted masks with detected wildfires. If the value of IoU for a pixel is higher than 0,5, it belongs to the class of interest.

As a loss function for real masks and predictions there was used a sum of binary cross entropy (BCE) and the value of Dice loss (DL):

Combinations of various loss functions for the learning of machine learning algorithms help to get higher quality of segmentation for the most modern tasks and data competitions [15].

Generated training and test sets were unbalanced, because the number of patches with objects is much less than the number of patches without objects. It can affect the learning process of developed CNN. To cope with this problem, a batch of images was formed by random selection of 8 patches with objects of «wildfire» class and 8 patches without them. This way of building patches for training showed the best segmentation results for our task. Also to control the learning process of U-ResNet34 there was formed a validation dataset which consist of 10% of all patches from training set taking into account the same policy of choice as in the case of building batches for training.

The training process on both databases finished after completing 100 epochs. On test process model used weights of CNN which were received on the training epoch with maximal value of DSC. Test results for U-ResNet34 on generated test datasets of the Resource database, which were enlarged by means of different techniques, were presented in Table 3.

**Table 3. Test results on test sets of the Resource database**

|  |  |  |  |
| --- | --- | --- | --- |
| **Metrics** | **Flips** | **Flips + SSR** | **Flips + SSR + HSV\_dist** |
| T | 0,32 | 0,50 | 0,50 |
|  | 0,371 | 0,465 | 0,436 |
| P | 0,348 | 0,456 | 0,431 |
| R | 0,396 | 0,475 | 0,442 |
| IoU | 0,874 | 0,871 | 0,870 |
| DSC | 0,782 | 0,812 | 0,794 |

According to results presented in Table 3, methods of data augmentation allow to improve results on test datasets for all metrics. In particular, -score reached 0,465 and the value of DSC was equal to 0,812. However, applying random chromatic distortion leds to slight degradation of quality of deep learning algorithm: developed model confused wildfires with clay areas. The example of an input image, true and predicted masks of the Resource database for U-ResNet34 is shown in Fig. 3.

**a) b) c)**

**Figure 3. Test results for U-ResNet34 on the Resource dataset: a) input image, b) true mask, c) predicted mask**

Test results for U-ResNet34 on generated test datasets of the Planet database, which were enlarged by means of different techniques, were presented in Table 4.

**Table 4. Test results on test sets of the Planet database**

|  |  |  |  |
| --- | --- | --- | --- |
| **Metrics** | **Flips** | **Flips + SSR** | **Flips + SSR + HSV\_dist** |
| T | 0,65 | 0,68 | 0,77 |
|  | 0,301 | 0,321 | 0,321 |
| P | 0,245 | 0,259 | 0,295 |
| R | 0,390 | 0,419 | 0,356 |
| IoU | 0,782 | 0,759 | 0,757 |
| DSC | 0,507 | 0,508 | 0,459 |

For images of the Planet database, test results were worse than result that were received for the Resource dataset: maximal -score reached 0,321 and the value of DSC was equal to 0,508. Also the quality of segmentation of correctly detected wildfires reduced: the value of IoU does not exceed 0.782. This is fact can be explained by the resizing affect of satellite images of the Planet database (12 m/pixel versus 10 m/pixel). Nevertheless, test results on the Planet dataset can be considered satisfactory. The example of an input image, true and predicted masks of the Planet database for U-ResNet34 is shown in Fig. 4.

**a) b) c)**

**Figure 4. Test results for U-ResNet34 on the Planet dataset: a) input image, b) true mask, c) predicted mask**

5. CONCLUSION

This article presents numerical experiments for developed deep learning algorithm: U-ResNet34. The training and testing process were performed on high-resolution aerial photos of the Resource and the Planet databases. To implement numerical experiments there were extracted smaller patches. Training and test sets were enlarged methods using various methods of data augmentation. According to test results for both datasets U-ResNet34 works satisfactorily: . For learning of model there was used supercomputer NVIDIA DGX-1.

ACKNOWLEDGMENT

The authors are also grateful to the AI-center of P.G. Demidov Yaroslavl State University for providing an access to NVIDIA DGX-1 supercomputer.

REFERENCES

[1] D.L. Martell, H. Sun. The impact of fire suppression, vegetation, and weather on the area burned by lightning-caused forest fires in Ontario. Canadian Journal of Forest Research-revue Canadienne De Recherche Forestiere, 2008, pp. 1547-1563.

[2] M. Flannigan, B. Amiro, K. Logan, B. Stocks, M. Wottom. Forest Fires and Climate Change in the 21st Century. Mitigation and Adaptation Strategies for Global Change, 2006, pp. 847-859.

[3] Wildfire Prevention Strategies. Web: https://web.archive.org/web/20081209105351/http://www.nwcg.gov/pms/docs/wfprevnttrat.pdf.

[4] Y. Goodfellow, Y. Bengio, A. Courville, Deep Learning, The MIT Press, 2016, 800 p.

[5] V. Khryashchev, L. Ivanovsky, V. Pavlov, A. Ostrovskaya and A. Rubtsov. Comparison of Different Convolutional Neural Network Architectures for Satellite Image Segmentation. In Proceedings of the 23rd Conference of Open Innovations Association FRUCT’23, Bologna, Italy, 2018, pp. 172-179.

[6] V. Khryashchev, V. Pavlov, A. Ostrovskaya, R. Larionov. Forest Areas Segmentation on Aerial Images by Deep Learning. 2019 IEEE East-West Design & Test Symposium (EWDTS), 2019, pp. 1-5.

[7] O. Ronneberger, P. Fischer, T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS, vol. 9351, 2015, pp. 234–341.

[8] A. Buslaev, S. Seferbekov, V. Iglovikov, A. Shvets. Fully Convolutional Network for Automatic Road Extraction from Satellite Imagery. CoRR, abs/1806.05182, 2018.

[9] DeepGlobe. CVPR 2018 – Satellite Challenge, Web: http://deepglobe.org.

[10] S. Bajpai. Implementing Convolutional Neural Network with Parallel Computing Using CUDA. International Journal of Innovative Science, Engineering & Technology, 2015, Vol. 2, Issue 11, pp. 517–520.

[11] R. Atienza, Advanced Deep Learning with Keras. Packt Publishing, UK, 2018, 368 p.

[12] Kingma, D., Ba, J. Adam: A Method for Stochastic Optimization. Web: https://arxiv.org/abs/1412.6980.

[13] S. Raschka Python Machine Learning, Packt Publishing Ltd., 2015, 454 p.

[14] J. VanderPlas. Python Data Science Handbook: Essential Tools for Working with Data. First Edition. O’Reilly Media, 2016, 541 p.

[15] Losses for Image Segmentation. Web: https://lars76.github.io/neural-networks/object-detection/losses-for-segmentation/.