# Wildfire Segmentation on Satellite Images using Deep Learning

Vladimir Khryashchev
*P.G. Demidov Yaroslavl State University*
Yaroslavl, Russia
v.khryashchev@uniyar.ac.ru

Roman Larionov
*P.G. Demidov Yaroslavl State University*
Yaroslavl, Russia
r.larionov1@uniyar.ac.ru

***Abstract*—Deep learning and convolutional neural network technologies are increasingly used in the problems of analysis, segmentation and recognition of objects in images. In this article a convolutional neural network for automated wildfire detection on high-resolution aerial photos is presented. Two databases of satellite RGB-images with different spatial resolution containing 1457 and 393 high-resolution images, respectively, were prepared for training and testing the neural network. Various techniques of data augmentation are used to enlarge training and test sets generated by data windowing. U-Net neural network with the ResNet34 as encoder was used in research. Neural network training was learning using the NVIDIA DGX-1 supercomputer. Adaptive moment estimation algorithm was used for optimization of training process. Special metrics, such as Sorensen-Dice coefficient, precision, recall, F1-score and IoU value allows to measure the quality of developed model. The developed algorithm can be successfully applied for early wildland fires detection in practical applications.**

***Keywords — image segmentation, satellite images, forest fire detection, deep learning***

## I. INTRODUCTION

A wildfire is a spontaneous fire that cannot be controlled, occurs outside a special focus, causes material damage and poses a danger to human life and health. Every year year thousands of fires occur for various reasons: seasonally dry periods, thunderstorms, volcanic ignitions. However, last years the human factor has become the main reason of causing irreversible forest fires.

Methods of wildfire detection and prevention have varied through 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.

There are more than 340 million hectares of forests are annually damaged by fire on the Earth. 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 [4, 5].

Remote sensing allows to track changes in large areas of the earth's surface, so the ground services can quickly respond to appeared fires [6, 7]. Computer vision methods are developoing rapidly, new papers about object detection are published `every day. But automated image segmentation has not reached to same quality as manual marking of high-resolution aerial photos [8]. Manual processing of satellite images shows the highest results, but it takes too much time and human resources. 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. This is the task of artificial intelligence. Most existing algorithms of solving segmentation tasks envolve using qof deep learning methods, especially convolutional neural networks (CNNs) [9].

This paper presents a developed convolutional neural network that can be used for wildland fires segmentation. Data preparation, CNN training, testing and discussion of the achieved results are presented. Our work continues research, which was presented in [10, 11].

## II. DATA PREPARATION

In our research, there were used the Planet dataset and the Resurs database from Resurs-P satellite. The Resurs 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].

978-1-7281-2572-5/20/$31.00 ©2020 IEEE

Satellite image segmentation envolves 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 corresponding binary mask of datasets have been sliced on parts of 256 × 256 pixels with the step of 128 by data windowing. The intersection of patches allows to fix the 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. Information about prepared patches of different databases is presented in Tables 1, 2. Examples of sliced images and masks of the Resurs and the Planet datasets are shown in Fig. 1, 2.

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:

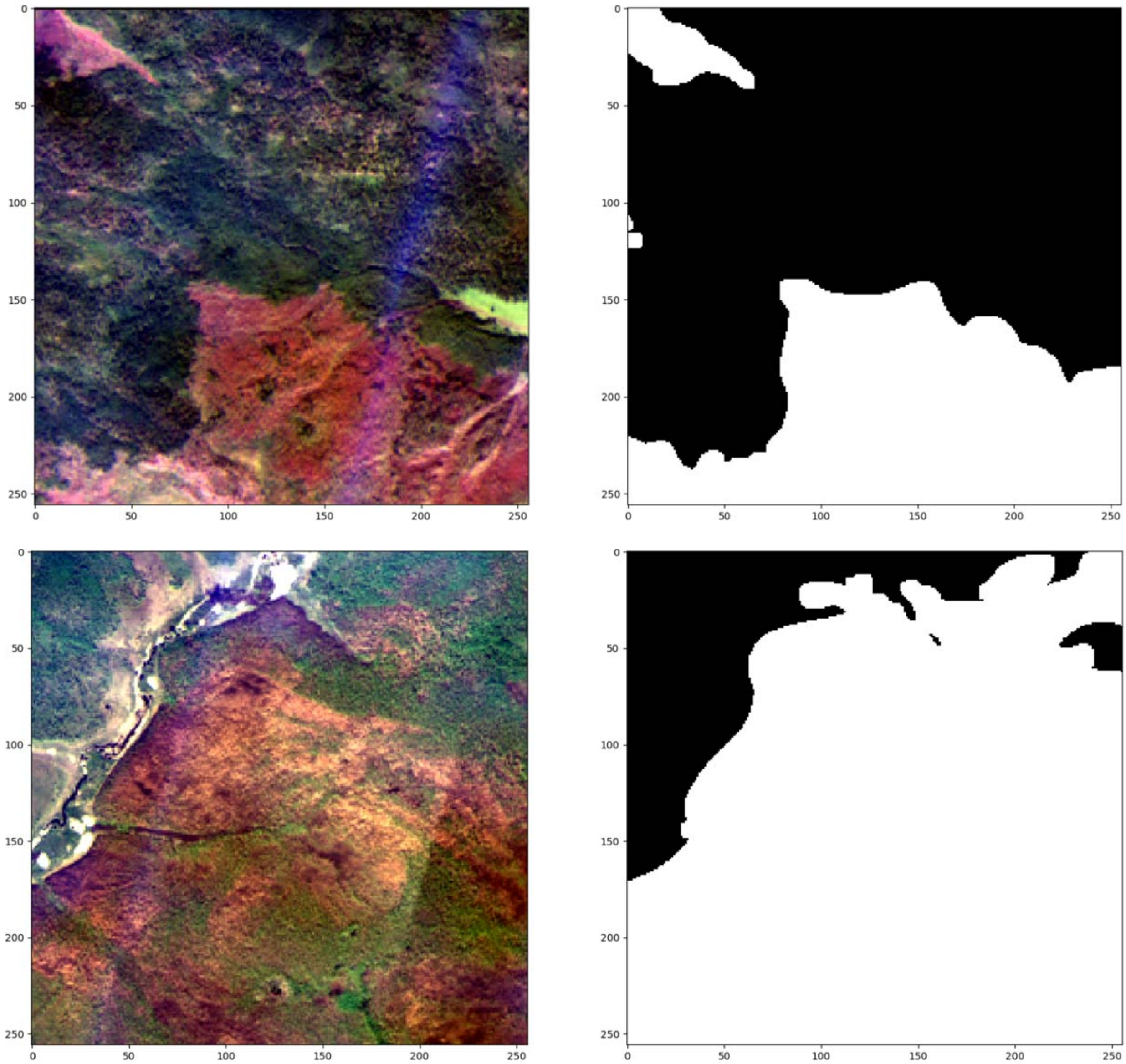
- rotations on 90˚, 180˚ and 270˚ and mirroring of patches. As a result, training and test sets were increased 8 times (flips);
- image shifts within 2% of image size, scaling on a coefficient from [1; 1,2] and rotations on small angles from [-15˚, + 15˚] (SSR).
- applying random chromatic distortion 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 (HSV_dist).

TABLE I. PREPARED PATCHES OF THE RESOURCE DATABASE

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

TABLE II. PREPARED PATCHES OF THE PLANET DATABASE

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



a) b)

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

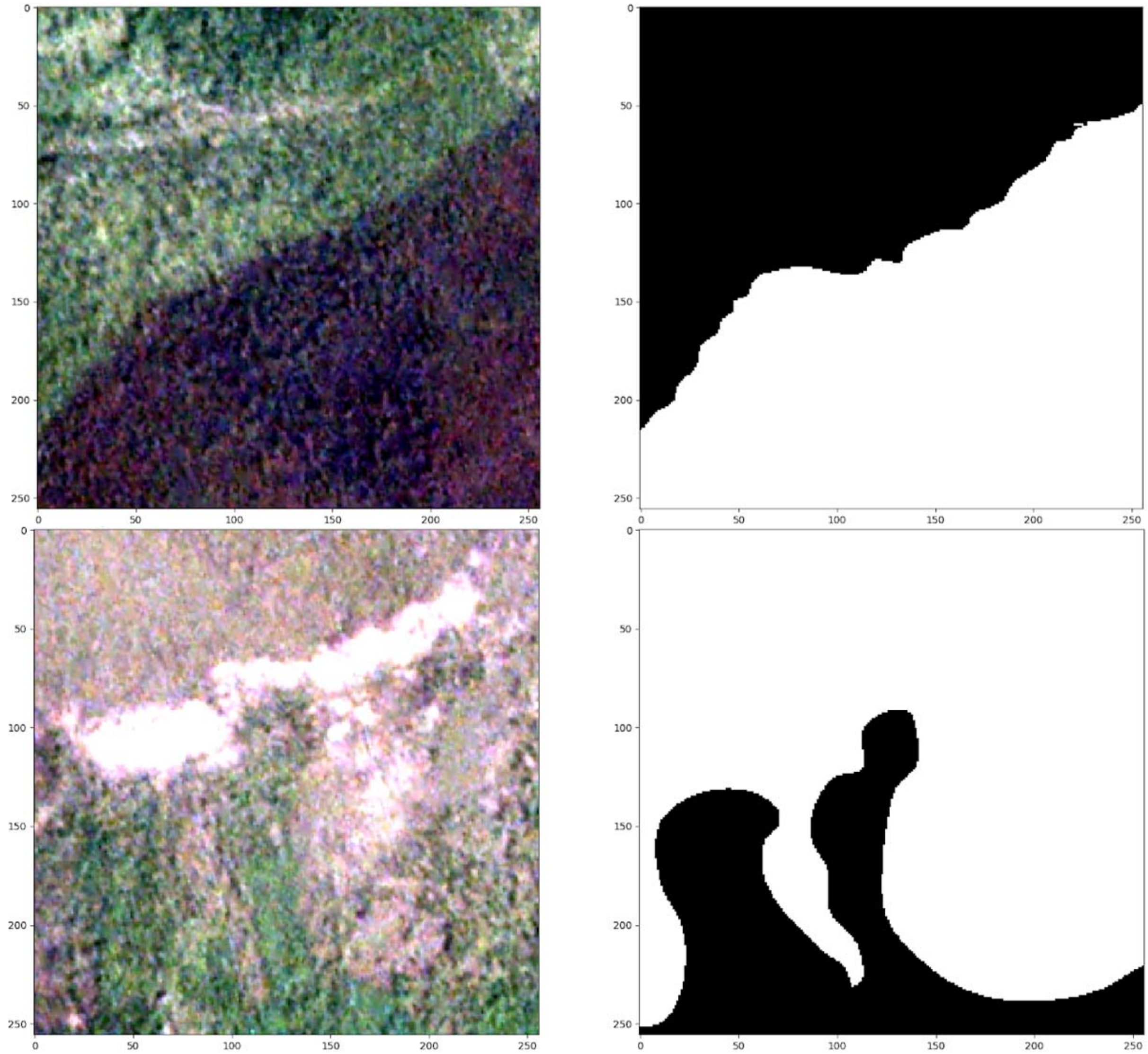



Fig. 2. Examples of a sliced image and a corresponding mask of the Planet dataset: a) prepared patches, b) true masks

## III. CONVOLUTIONAL NEURAL NETWORK

The CNN, which was used for wildfire segmentation on satellite images in our research is based on famous architecture called U-Net. This CNN was originally developed for medical images segmentation. Its classical structure is described in [12]. For the first time U-Net showed its efficiency for the task of medical image segmentation. Later this model has shown great results in geophysic images segmentation with small amount of training data. 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 is symmetrical part restorating the segmentation mask. The main advantages of the network are the achievement of high results with a small number of training images and a short duration of training process [13, 14].

In our research there was developed UNet-like architecture – U-ResNet34 based on models from paper [12]. 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 [13, 14]. This algorithm shows excellent results of detecting roads on satellite images of DeepGlobe database [16].

Adaptive moment estimation (Adam) was used for optimization of training process. It's the modification of Alagard method and combines the idea of accumulation of movement and the idea of weaker updating of weights for typical features [17].

## IV. NUMERICAL RESULTS

Special metrics are used to evaluate the quality of segmentation algorithm. In our research there was used Sorensen-Dice coefficient (DSC), precision ($P$), recall ($R$), $F_1$-score ($F_1$) and $IoU$ coefficient evaluated for true positive examples. The value of $F_1$-score was used as the main metric of measuring the quality of developed CNN. $F_1$-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 [18, 19]. $F_1$-score is calculated by the following formula:

$$F_1 = 2PR/(P + R).$$

During the training, there was chosen a threshold value ($T$) 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 $X$ and predictions $Y$ there was used a sum of binary cross entropy (BCE) and the value of Dice loss (DL):

$$Loss = BCE(X,Y) + DL(X,Y),$$

$$BCE(X,Y) = -\Sigma_{x,y}(x\log(y) + (1 - x)\log(1 - y)),$$

$$DL(X,Y) = 1 – 2|X \cap Y| / (|X| + |Y|)$$

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 [20].

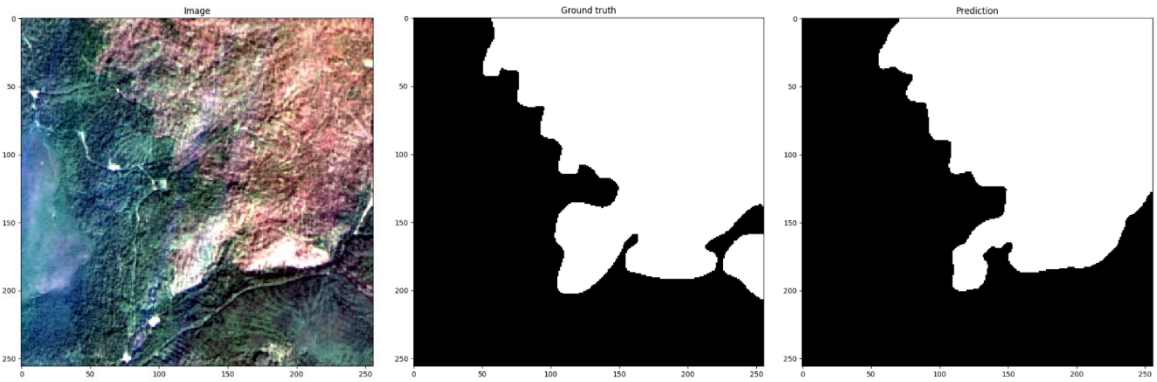
Generated training and test sets were unbalanced, because the number of patches with objects is much less than the number of patches without objects. It can affect the learning process of developed CNN. To cope with this problem, a batch of images was formed by random selection of 8 patches with objects 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 Resurs database, which were enlarged by means of different techniques, were presented in Table 3.

According to results presented in Table 3, methods of data augmentation allow to improve results on test datasets for all metrics. In particular, $F_1$-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 Resurs database for U-ResNet34 is shown in Fig. 3.

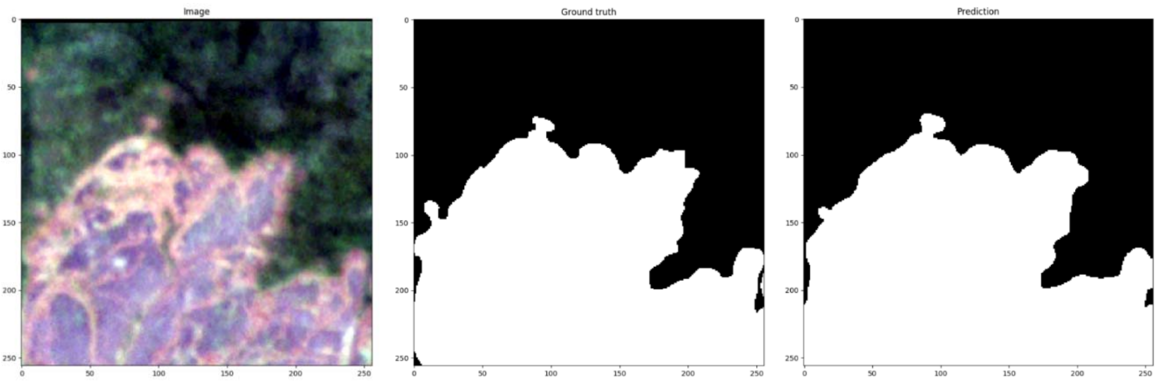
Test results for U-ResNet34 on generated test datasets of the Planet database, which were enlarged by means of different techniques, were presented in Table 4.

For images of the Planet database, test results were worse than result that were received for the Resurs dataset: maximal $F_1$-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)

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




a) b) c)

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

TABLE III. TEST RESULTS ON TEST SETS OF THE RESURS DATABASE

| Metrics | Flips | Flips + SSR | Flips + SSR + HSV_dist |
|---|---|---|---|
| T | 0,32 | 0,50 | 0,50 |
| F1 | 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 |

TABLE IV. TEST RESULTS ON TEST SETS OF THE PLANET DATABASE

| Metrics | Flips | Flips + SSR | Flips + SSR + HSV_dist |
|---|---|---|---|
| T | 0,65 | 0,68 | 0,77 |
| F1 | 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 |

## V. 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 Resurs 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.

## ACKNOWLEDGMENT

This work was supported by a grant from the P.G. Demidov Yaroslavl State University №ОП-2Г-09-2019

## 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] H. Huang, C. Li, X. Wang, X. Zhou, “Integration of multi-resource remotely sensed data and allometric models for forest aboveground biomass estimation in China”, Remote Sensing of Environment, 2019. Vol. 221. Pp. 225-234.

[5] O. Ghorbanzadeh, T. Blaschke, K. Gholamnia, D. Tiede, “Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection”, Remote Sensing, 2019. Vol 11(2). Pp. 1-21.

[6] Y. Yu, K. Zhang, L. Yang, D. Zhang, “Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN”, Computers and Electronics in Agriculture, 2019. Vol. 163. Pp. 1-9.

[7] T. Zhao, Y. Yang, H. Niu, D. Wang, “Comparing U-Net convolutional network with mask R-CNN in the performances of pomegranate tree canopy segmentation”, In Proceedings of SPIE - The International Society for Optical Engineering, 2018. Vol. 10780. Pp. 1-9.

[8] J. Yang, Y. He, J. Caspersen, “A multi-band watershed segmentation method for individual tree crown delineation from high resolution multispectral aerial image”, Geoscience and Remote Sensing Symposium (IGARSS), 2014. Vol. 2014. Pp. 1588-1591

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

[10] 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.

[11] 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.

[12] 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.

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

[14] N. Khalid, J. A. Hamid, Z. Latif A., “Accuracy assessment of tree crown detection using local maxima and multi-resolution segmentation”, In IOP Conference Series: Earth and Environmental Science. 2014. Vol. 18(1). Pp. 1-6.

[15] J. Yang, Y. He, J. Caspersen, “A multi-band watershed segmentation method for individual tree crown delineation from high resolution multispectral aerial image”, Geoscience and Remote Sensing Symposium (IGARSS), 2014. Vol. 2014. Pp. 1588-1591.

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

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

[18] S. Raschka, “Python Machine Learning”, Packt Publishing Ltd., 2015, 454 p.

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

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