# U-NET MODEL FOR LOGGING DETECTION BASED ON THE SENTINEL-1 AND SENTINEL-2 DATA

*Leonid Shumilo[1,2], Nataliia Kussul[1,2], Mykola Lavreniuk[1]*

[1] Space Research Institute NASU-SSAU, Kyiv, Ukraine
[2] National Technical University of Ukraine "Igor Sikorsky Kiev Polytechnic Institute", Kyiv, Ukraine

## ABSTRACT

Illegal logging in Ukraine is a big problem that negatively affects both environmental and socio-economic indicators of the country. The main reason for this problem is the lack of independent control over the forest industry. Lack of control, in turn, makes it possible to provide inaccurate information about the permitted logging and to hide the fact of logging. The solution to this problem is the use of modern approaches of Remote Sensing and deep learning to implement mechanisms for forestry monitoring and logging detection based on the satellite data. Most researches on satellite-based logging detection technology are based on the optical satellite missions. However, for countries with temperate and cold climates, the use of such approaches is problematic in winter and autumn due to the lack of vegetative biomass and the high percentage of clouds and snow in satellite images. In this study, we assessed a methodology for detecting logging based on optical and radar images of Copernicus satellite missions, namely Sentinel-1 and 2. The obtained results show that when using this approach, it is possible to monitor and detect logging with high accuracy both in summer and in winter with the frequency of data updates once a week. The basis of this methodology is a convolutional neural network with U-Net architecture, which input is a stack of optical and radar images in summer and spring, and works on radar images only in winter and autumn.

***Index Terms*****—** Deep Learning, Sentinel-1, U-Net, logging detection.

## 1. INTRODUCTION

Satellite data have long been the basis for forestry monitoring in the world. One of the most famous global projects is Global Forest Watch [1]. This project uses satellite data to build forest maps for the world, as well as deforestation alerts for tropical forests with a significant problem of their extinction. Different approaches of land cover classification are used to solve the problem of forest map production. The most common method of machine learning for this is the use of decision trees and random forest. Such classic pixel-based approaches in the presence of high-quality training data and long time-series of satellite data give a stable and high-quality result [2]. However, pixel-based approaches have the problem of noise, the removal of which is an important part of post-processing of such maps. Also, the presence of noise complicates the process of monitoring logging in near real time, because in this case the time series is limited to a small number of dates and the size of objects of interest is also small, so these objects can be confused with noise. At the same time, deep learning classification approaches are becoming increasingly popular in the problems of land cover classification. The use of convolutional neural networks, despite the relative difficulty in choosing qualitative hyperparameters, give a better result than pixel-based methods [3]. There are adapted to satellite data architectures of recurrent neural networks such as LSTM [4], which have advantages in work with data presented in time series. Image segmentation approaches from a computer vision are actively used to solve the problems of object detection, to determine agriculture fields boundaries, to build maps of infrastructure to build maps land cover and etc. The main feature of such methods is the ability to use them on a very small number of images and get a good result with insignificant amount of noise, even on a single satellite image. The most popular such architecture is U-Net, which is widely used for segmentation of satellite images [5].

The greatest development of forest monitoring technologies occurred in the opening of access to Landsat-8 data and the launch of satellite missions of the Copernicus programme, Sentinel-2 and Sentinel-1. Today, Sentinel-2 is the most popular open source of satellite information in country or regional logging detection projects based on deep learning approaches [6]. The reason for this is that on optical data, objects such as logging are very visible, especially if take into account high spatial resolution of such data 10 meters. However, the problem with this approach is that optical images are affected by clouds, leading to a lack of information for a significant part of the time series, and in the autumn and winter months they are poorly applicable due to lack of active vegetative biomass in forests and snow cover. These problems can be solved by using radar data along with optical. Radar data in this way are used for the most part in long time series, requires complex processing workflow and provides a result with more noise [7]. However, on cloudy days, as well as in autumn and winter, it is the only source of information.

## 2. STUDY AREA AND MATERIALS

### 2.1. Geography

Ukraine is the biggest country fully located in Europe. The overall area of Ukraine is 603,548 km$^2$ and the biggest part of territory (68%) used for agricultural purpose. As of 2019, the area of forests in Ukraine is 95,739 km$^2$ by the official statistics of State Agency of forest resources of Ukraine and the actual forest cover of the territory is 15,9%. The 73% of Ukrainian forests are on governmental balance. According to official statistics, the volume of illegal logging in 2019 was 118,000 m$^3$, while the total volume of felled wood is 15,6 km$^3$. An important problem in Ukraine is the lack of control over felling, so permission to carry out legal felling is often obtained illegally by providing inaccurate information.

The study area in this research is Kyiv region, which has 28121 km2 of forest area and 22% of forest cover. On the territory of Kyiv region there are deciduous, coniferous and mixed forests.

### 2.2. Satellite and in-situ data

During the experiment, we used 2 images of Sentinel-1 with relative orbit 87 for 2020.08.28 and 2020.12.08 and 1 image of Sentinel-2 35UQS for 2020.08.30. Optical image was processed with use of sen-2-core software and processing include atmospheric correction, radiometric correction and cloud masking. Sentinel-1 data were processed with use of SNAP software and processing graph that include refined lee filtering, calibration and radiometric correction. Figure 1 show the comparison of visual view of logging on Sentinel-1 VV and VH 2-band image and optical Sentinel-2 RGB image.

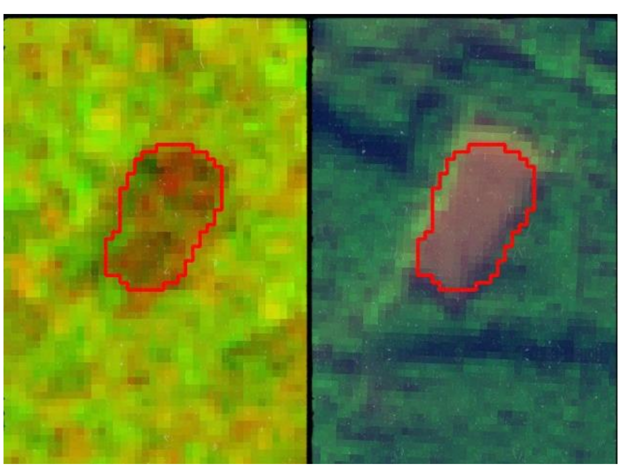

Figure 1. Logging on Sentinel-1 (left) and Sentinel-2 (right) images

Ground truth samples were produced based on the photointerpretation of satellite images and 3507 polygons of logging were used in the experiment. The testing area was separated from training area geographically and include 2011 polygons from ground truth set. Figure 2 shows the distribution of logging polygons on the Sentinel-2 image.

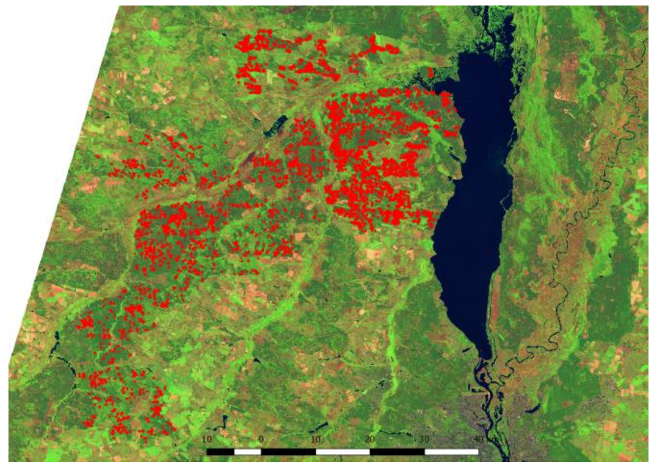

Figure 2. Ground truth samples (polygons in red) on the Sentinel-2 image

## 3. METHODOLOGY

For our experiments we took one of the most appropriate and accurate deep learning architectures for the semantic segmentation tasks – U-net model. The architecture of our U-net model traditionally consists of convolutional part and deconvolutional that connected with each other using concatenation operation (Fig. 3). In convolutional part we utilized several blocks that consists of consequently applied convolutional operation with ReLU activation function, batch normalization layer for better network converges within the training phase. After each two convolutional blocks maxpooling layer has been used for decreasing the width and the height of the tensor with 2x factor.

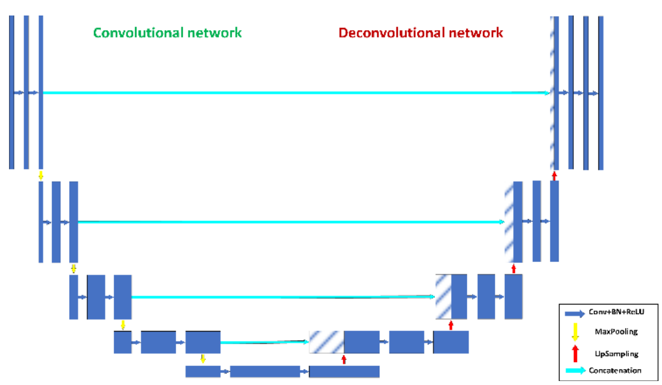


Figure 3. The architecture of the U-net model for semantic segmentation tasks

Traditionally, for training multi-layer perceptron and the deep learning models cross-entropy (CE) has been chosen as the loss function [8] or weighted CE

$$CE = -\sum_{k=1}^{N}\sum_{c=1}^{C}\alpha_c y_c \log\left(p_c\right), \qquad (1)$$

where C – number of classes, N – number of samples, y – target vector, in our case we choose one-hot coding for it, p – output of the last layer of the neural network, $\alpha_c$ – coefficient for control the class impact, basically $\alpha_c = 1$ or $\alpha_c = \frac{N_c}{N}$.

In this paper we wanted not only to train the U-net model with CE, but also to conduct the experiments with novel and more specific loss functions that are more suitable for tasks with imbalanced training data. In the paper [9] authors proposed focal loss (FL) for the object detection and localization task

$$FL = -\sum_{k=1}^{N}\sum_{c=1}^{C}\alpha y_c \left(1-p_c\right)^{\gamma}\log\left(p_c\right), \qquad (2)$$

where $\alpha, \gamma$ – coefficients, traditionally $\alpha = 0.25, \gamma = 2$.

The main advantage of this loss function is decreasing an impact of the of the large number of negative samples compare to the small number of positive samples. However, focal loss is design for object detection task it could be adopted and suitable for solving tree logging problem using satellite data regarding to imbalance samples distribution.

Specially for accuracy assessment in semantic segmentation tasks intersection over union (IoU) coefficient is used

$$IoU(Y,P) = \frac{Y \cap P}{Y \cup P}, \qquad (3)$$

where Y – binary target imagery, P – output probability for certain class.

Thus, for semantic segmentation tasks often it is more effective to train neural network to increase the accuracy coefficient (decrease Jaccard loss) [10, 11] rather than to decrease common CE

$$Jaccard = 1 - \frac{1}{N_c}\sum_{c=1}^{C}\frac{Y_c \cap P_c}{Y_c \cup P_c}. \qquad (4)$$

Another specific for semantic segmentation tasks with imbalanced training samples distribution is the Dice loss [12] that could be calculated in two different ways

$$Dice = 1 - \frac{2}{N_c}\sum_{c=1}^{C}\frac{\left|Y_c \cap P_c\right|}{\left|Y_c\right| + \left|P_c\right|}. \qquad (5.1)$$

$$Dice = 1 - \frac{2}{N_c}\frac{\sum_{c=1}^{C}\alpha_c\left|Y_c \cap P_c\right|}{\sum_{c=1}^{C}\alpha_c\left(\left|Y_c\right| + \left|P_c\right|\right)}, \qquad (5.2)$$

where $\alpha_c$ – coefficient for control the class impact, basically $\alpha_c = 1$ or $\alpha_c = \frac{N_c}{N}$.

For model accuracy assessment we have chosen IoU, User Accuracy (UA), Producer Accuracy (PA) and combination of UA and PA, namely F1-score. All these metrics have been calculated for logging class only. Such decision is mainly due to large imbalance not only in training data, but also in test data, which lead to very high accuracies for another class and their changes are not sufficient and interesting in solving our problem.

## 4. RESULTS

The results for all experiments with different input data and different loss functions are shown in Table 1 and the result images are shown in Figure 4. The results received within this study confirm common thought that optical data is more appropriate for forest monitoring and particularly logging detection (F1-score and IoU metric are higher for S2 summer image comparing to S1 summer image). At the same time fusion of Sentinel-1 and Sentinel-2 received the most accurate results with F1-score = 70.2 and IoU = 0.54. However, the gain from Sentinel-1 is not significant (less than 2 points for each metrics).

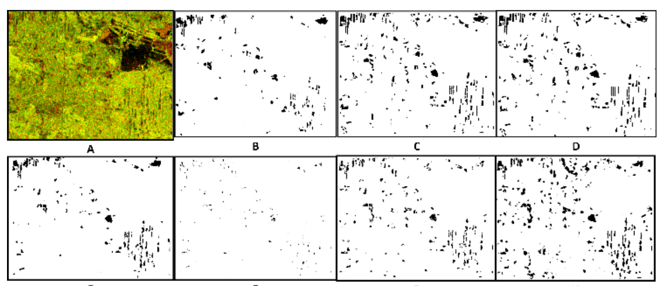


Figure 4. Results of the different U-net models: A) S1 summer image, B) S1 summer with CE, C) S2 summer with CE, D) S1+S2 summer with CE, E) S1 winter with CE, F) S1 winter with Jaccard, G) S1 winter with Dice and H) S1 winter with Focal loss.

Interesting fact that was found it is not only that Sentinel-1 is the only source of available satellite information for logging detection within the winter season, but also that obtained results for winter is much better than results based on the Sentinel-1 for the summer season.

Comparison of different specific for semantic segmentation tasks losses shows that despite the fact that Jaccard loss provides similar results as traditional CE loss withing the UA metrics, all other metrics were below the appropriate results. Quite interesting was the fact that Dice loss provides better results compared to Jaccard loss and also all metrics is balanced.

Focal loss provides quite lower results compared to Dice loss (F1-score = 39.2 and IoU = 0.24), but the balance of the metrics is not consistent as it is within the utilizing of the Dice loss.

Table 1. Comparison of user accuracy (UA), producer accuracy (PA), F1-score and intersection over union (IoU) for the different input and different loss functions

| | Summer S1 Cross-entropy | Summer S2 Cross-entropy | Summer S1+S2 Cross-entropy | Winter S1 Cross-entropy | Winter S1 Jaccard | Winter S1 Dice | Winter S1 Focal loss |
|---|---|---|---|---|---|---|---|
| UA | 64.8 | 64.5 | 66.4 | 69.8 | 64.0 | 47.6 | 29.2 |
| PA | 40.3 | 73.0 | 74.3 | 45.5 | 10.7 | 44.4 | 59.6 |
| F1 | 49.7 | 68.5 | 70.2 | 55.1 | 18.3 | 45.9 | 39.2 |
| IoU | 0.33 | 0.52 | 0.54 | 0.38 | 0.10 | 0.30 | 0.24 |

## 5. DISCUSSION AND CONCLUSIONS

This study compared the possibilities of using satellite optical and radar data in the problem of logging detection, using a convolutional neural network with the architect U-Net. Experiments have shown that the best results can be obtained by combining radar and optical images. However, also an important conclusion of this study is that radar data give sufficient accuracy in solving this problem in the absence of optical images. This makes it possible to use deep learning approaches for logging detection on cloudy days and in autumn and winter, when optical data are absence.

## ACKNOWLEDGMENT

The authors acknowledge the funding received by the National Research Foundation of Ukraine from the state budget 2020.02/0292 "Deep learning methods and models for applied problems of satellite monitoring" (NRFU Competition "Support for research of leading and young scientists") and was supported by European Commission “Horizon 2020 Program” that funded ERA-PLANET/GEOEssential (GrantAgreement no. 689443).

## REFERENCES

[1] Watch, G. F. (2002). Global forest watch. World Resources Institute, Washington, DC Available from http://www. globalforestwatch. org (accessed March 2002).

[2] Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., ... & Townshend, J. (2013). High-resolution global maps of 21st-century forest cover change. science, 342(6160), 850-853.

[3] Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. IEEE Geoscience and Remote Sensing Letters, 14(5), 778-782.

[4] Rußwurm, M., & Körner, M. (2017). MULTI-TEMPORAL LAND COVER CLASSIFICATION WITH LONG SHORT-TERM MEMORY NEURAL NETWORKS. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 42.

[5] Mboga, N., Grippa, T., Georganos, S., Vanhuysse, S., Smets, B., Dewitte, O., ... & Lennert, M. (2020). Fully convolutional networks for land cover classification from historical panchromatic aerial photographs. ISPRS Journal of Photogrammetry and Remote Sensing, 167, 385-395.

[6] Isaienkov, K., Yushchuk, M., Khramtsov, V., & Seliverstov, O. (2020). Deep Learning for Regular Change Detection in Ukrainian Forest Ecosystem with Sentinel-2. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 364-376.

[7] Dostálová, A., Wagner, W., Milenković, M., & Hollaus, M. (2018). Annual seasonality in Sentinel-1 signal for forest mapping and forest type classification. International Journal of Remote Sensing, 39(21), 7738-7760.

[8] Bengio, Yoshua, Ian Goodfellow, and Aaron Courville. *Deep learning*. Vol. 1. Massachusetts, USA: MIT press, 2017.

[9] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2980-2988).

[10] Jadon, S. (2020, October). A survey of loss functions for semantic segmentation. In *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)* (pp. 1-7).

[11] Bertels, J., Eelbode, T., Berman, M., Vandermeulen, D., Maes, F., Bisschops, R., & Blaschko, M. B. (2019, October). Optimizing the dice score and jaccard index for medical image segmentation: Theory and practice. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 92-100). Springer, Cham.

[12] Fidon, L., Li, W., Garcia-Peraza-Herrera, L. C., Ekanayake, J., Kitchen, N., Ourselin, S., & Vercauteren, T. (2017, September). Generalised wasserstein dice score for imbalanced multi-class segmentation using holistic convolutional networks. In *International MICCAI Brainlesion Workshop* (pp. 64-76). Springer, Cham.