Advanced Method of Land Cover Classification Based on High Spatial Resolution Data and Convolutional Neural Network

Andrii Shelestov[1], Bohdan Yailymov[2], Hanna Yailymova[1], Leonid Shumilo[3], Mykola Lavreniuk[2], Alla Lavreniuk[1], Sergiy Sylantyev[2] and Nataliia Kussul[1]
[1]Institute of Physics and Technology, NTUU "Igor Sikorsky Kyiv Polytechnic Institute", 37 Peremohy avenue, Kyiv, Ukraine
[2]Department of Space Information Technologies and System, Space Research Institute NAS Ukraine an SSA Ukraine, 40 Glushkov avenue, Kyiv, Ukraine
[3]Department of Geography, The University of Maryland, College Park, Maryland, USA
{andrii.shelestov, yailymov, anna.yailymova, shumilo.leonid, alla.lavrenyuk, sylantyev, nataliia.kussul}@gmail.com, nick_93@ukr.ne

Keywords: Convolution Neural Network, Probability Classification, Land Cover Map, Urban Atlas, Smart City.

Abstract: Based on modern satellite products Planet with high spatial resolution 3 meters, authors of this paper improved the neural network methodology for constructing land cover classification maps based on satellite data of high spatial resolution using the latest architectures of convolutional neural networks. The process of information features formation for types of land cover is described and the method of land cover type classification on the basis of satellite data of high spatial resolution is improved. A method for filtering artificial objects and other types of land cover using a probabilistic channel is proposed, and a convolutional neural network architecture to classify high-resolution spatial satellite data is developed. The problem of building density maps for the quarters of the city atlas construction is solved and the metrics for estimating the accuracy of classification map construction methods are analyzed. This will make it possible to obtain high-precision building maps to calculate the building area by functional segments of the Urban Atlas and monitor the development of the city in time. This will make it possible to create the first geospatial analogue of the product Copernicus Urban Atlas for Kyiv using high spatial resolution data. This Urban Atlas will be the first such product in Ukraine, which can be further extended to other cities in Ukraine. As a further development, the authors plan to create a methodology for combining satellite and in-situ air quality monitoring data in the city based on the developed Urban Atlas, which will provide high-precision layers of PM_{10} and $PM_{2.5}$ concentrations with high spatial and temporal resolution of Ukraine.

1 INTRODUCTION

The image of remote sensing data is a matrix of "pixels", which are the smallest unit of the image and contain the values of the measured spectral reflectance. Traditionally, pixel-based approaches are used to classify images, as a result of which each pixel of the image is mapped to a class according to its spectral properties [1]. Object characteristics include spectral characteristics, shape, size, texture, and context properties. Signs usually determine the upper and lower limits of the range of measured characteristics of objects. Image objects within certain contours belong to a certain class, and those outside them belong to other classes [2]. An important task is to identify the most relevant features for each class in order to effectively perform the classification of images with high accuracy [3]. The following features are most often used in the literature for the classification of urban areas: brightness, average value of the reflection spectrum and standard deviation of the spectra [4]. According to the average value and the standard deviation of the features, the threshold values of the selected features are determined, which are used to assign a class label to each pixel [32]. To understand the relevance of the research problem and answer key issues that researchers face when classifying land cover using high-resolution data, we analyzed existing approaches related to the pixel and object-oriented approach using convolutional neural networks

(CNN), which gave us the opportunity to improve our own algorithm for classifying urban areas on Sentinel-2 satellite data with a spatial resolution of 10 meters, as well as to develop a new algorithm for classifying satellite data with a high spatial resolution (for example, Planet data with a spatial resolution of 3 meters).

Within the object-oriented approach, the smallest spatial unit for mapping is geospatial objects, which look like vector polygons extracted from high-resolution remote sensing images using segmentation techniques. These internally homogeneous geographical objects are used as the basic units of the classification map. Conditionally multi-scale segmentation is used to select homogeneous image objects and create appropriate vector polygons [5].

The purpose of this article is to consider the main problems of monitoring and classification of land cover in Kyiv on the basis of our proposed neural network algorithm, which provides an opportunity to update annually for city's land cover maps with high spatial resolution, which is crucial information in management decisions when planning the cities development in the long run.

In recent years, segmentation methods have been proposed with further classification based on CNN to extract the boundary information of objects using self-learning functions from images [6]. A study by Dong, Wu, Luo, Sun, and Xia (2019) proposed a method for determining geospatial objects based on CNN [7]. The method consists of the following stages.

Initially, road and river landfills on the historical map of the earth's crust are used to zone the target image into several subregions. Subsequent object classification of subregions can be performed in parallel mode. In the second stage, a probability map is built for each subregion, using a modified convolutional network VGG16 [8]. This network uses five convolution layers, the output of which is combined by the upsampling method [7]. In the third stage, a vector probability map of object boundaries is built. The subregion images are then extracted from the landfill boundary of each geospatial object. The results of all subregions are combined to form a structured map of geospatial vector classification.

After determining the shape of geospatial objects for each of them the procedure of extraction of features is performed. Traditionally, each class of geospatial objects has a set of features that are determined on the basis of satellite data. Let's focus on the following image characteristics: spectrum, shape and texture. The spectral characteristics the average pixel brightness for the selected object, the standard deviation and maximum differences of the spectral signals, the normalized differential vegetation index (NDVI) and the normalized differential water index (NDWI) are usually used. These indices are calculated based on satellite data with high spatial resolution (one value per pixel), and the average values of the corresponding pixels that make up the geospatial object are used as features of the object itself. Indicators of the shape of geospatial objects are the size of the object, the ratio of length to width, pixel index of shape and the overall index of shape [9]. The textural features of geospatial objects use the index PanTex of the presence of buildings [10], and the degree of matrices of the gray-level co-occurrence matrices (GLGM) [11]. Based on the data of the digital terrain model, you can also calculate the terrain characteristics, including the average topographic height and steepness for each pixel.

The corresponding features of the object are calculated as the average values of the pixels that are part of each geospatial object. In Table 1 lists the features of the classes used in deep learning algorithms are presented.

Table 1: The List of Features Used in the Classification by Deep Learning.

Spectrum characteristics	Texture features	Topographic features
The average value of the spectrum signals in different satellite channels	Homogeneity	Height
Standard deviation of spectral signals in different satellite channels	Contrast	Incline
Brightness of spectral signals	Unlikeness	
Maximum differences of spectral signals	Entropy	
Normalized vegetation index (NDVI)	Correlation	
Normalized water index (NDWI)		

In Figure 1 presents the NDVI values for selected types of land cover and different geospatial objects. You can use additional geospatial data from different sources to obtain additional informational features.

Different land cover classification methods based on high spatial resolution data can be very useful for different land classification problems under different urban agglomerations, making it possible to obtain an accurate geospatial vector urban atlas like the Copernicus Urban Atlas. The use of satellite data with a spatial resolution of 10 meters does not make it possible to accurately identify artificial objects in suburban areas.

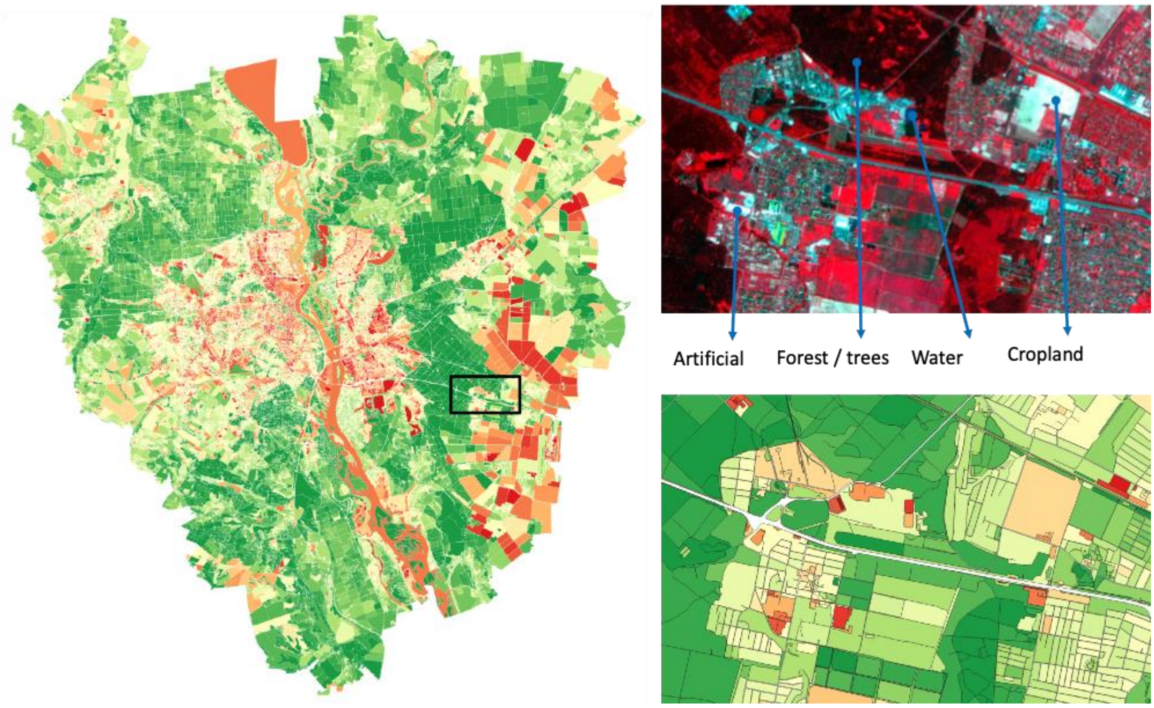

Figure 1: Examples of NDVI values for selected types of land cover and various geospatial objects of the Kyiv city and its environs based on Planet data.

In recent years, the Object-based image land cover classification has attracted considerable attention based on objects with high spatial resolution data and a convolutional neural network for mapping the earth's surface using remote sensing images. Numerous studies over the past decade have examined a wide range of issues related to land mapping [1] – [8]. However, most of these studies focused on agricultural land and crop types. In our case, it is important to accurately identify artificial objects, the identification of which on the satellite data of high spatial resolution may cause problems with shadows from large buildings, which belong by the neural network to the incorrect class. This is our research - to improve existing and develop our own neural network algorithm for the classification of urban areas.

2 METHOD

2.1 Study Area

Mapping urban land use and land cover is a fundamental task of urban planning and management. Very high resolution (VHR) satellite images, such as images from QuickBird, IKONOS, GeoEye, WorldView-2/3/4 and GaoFen-2, have shown a great advantage in urban land monitoring conditions due to high spatial detailing, compared to free data. In recent years, many studies have been conducted in the world on the classification of urban land use and soil cover based on VHR images [12], [34], [35], [37]. High-resolution classification methods can generally be divided into two classes - pixel-based methods and object-based methods [13]. The first defines the classes for individual pixels mainly on the basis of spectral information. In VHR images, pixel-by-pixel methods can be problematic because the spectral values of individual pixels on the earth's surface do not reflect the characteristics of the object. Object-oriented methods have been proposed to solve this problem. Unlike pixel-by-pixel methods, object-oriented methods combine adjacent pixels into objects using image segmentation, such as Multi-Resolution, Mean-Shift or Quadtree-Seg, and the objects are considered as units of classification [14]. Object-oriented classification methods greatly help to reduce problems with pixel-by-pixel approaches, while the quality of image segmentation significantly affects the accuracy of classification [13]. In addition, analysts must first randomly select a set of object features or design representative features as input to the classification using design techniques methods. The effectiveness of the selected features also affects the accuracy of the classification. Objects selected in a complex urban area are unlikely to be representative of all types of land cover. To solve this problem, it is necessary to provide automatic selection of features on images from remote sensing satellites instead of using manually selected features.

2.2 Classification Procedures

Artificial intelligence methods for image recognition and computer vision, such as K-means, neural networks (NN), support vector machines (SVMs), and random forests (RF), are not effective ways to classify Earth remote sensing images. A study by Fu, Zhao, Li, and Shi (2013) proposed for those purposes the idea of deep learning [14]. Compared to traditional machine learning methods such as NN, SVM and RF, deep learning models provide automatic feature extraction from large data sets. They learn to remove key features in the modeling process; thus, do not require prior selection of signs. In computer vision communities, deep learning, including convolutional neural networks (CNNs), has been used successfully to categorize images, identify goals, and understand scenes [15]. As a rule, typical CNNs include convolutional and aggregation layers, a nonlinearity activation function followed by fully bound layers as classifiers. CNN's first application of remote sensing was the allocation of road networks and buildings by Mnih and Hinton (2010) [16]. CNN has recently been used for pixel-by-pixel semantic labeling (or classification) of high-resolution remote sensing images. Mnih and Hinton (2010) proposed the CNN architecture for aerial imaging. Wang et al. (2015) used a three-layer CNN structure and a finite state machine to identify the road network [17]. Hu et al. (2015) used a pre-trained CNN network to classify different scenes in high-resolution remote sensing images [18]. Längkvist and others (2016) used CNN for the pixel-by-pixel classification of 0.5 m aerial photographs in natural colors into five classes of land cover, including vegetation, land, roads, buildings, and water, using the Digital Surface Model (DSM) [19]. Maltezos (2016) used CNN to identify buildings based on ortho-photo images using object height information [20]. Pan and Zhao (2017) proposed an extended CNN model for land cover classification based on 4-channel GaoFen-2 satellite images of rural areas [21]. Long at al. (2015) proposed a fully convolutional neural network (FCN) [22]. In FCN, fully connected layers in CNN are replaced by layers with upward convolution and merged with a shallow layer. Because the standard CNN model is based on the "image label" principle, the FCN's "start-to-end" labeling mode is more suitable for pixel-based image classification, and assigning a specific class to each pixel. The structure of the FCN has also shown great potential in the classification of remote sensing images. Sherrah (2016) proposed the FCN structure without decreasing sampling for semantic marking based on reference data from the International Society for Photogrammetry and Remote Sensing (ISPRS) in Vaihingen and Potsdam, which are publicly available aerial photographs in conventional colors with spatial resolution 9 cm and accompanied by DSMs derived from Lidar [23]. Based on the same data sets, implemented a multi-scale FCN [24]. The ensemble FCN model is offered [25]. Maggiori at al. (2017) compared CNN and FCN models for aerial photography and presented a multilayer perceptron (MLP) structure based on the FCN model for large-scale aerial photography classification [26]. Our review of the literature shows that most of the existing studies, with exception of [24], were conducted on basis of test sets ISPRS. Although these publicly available datasets provide images and reference sources for more efficient modeling and model comparisons, lower-spatial satellite imagery is often used for operational land classification and land use in large urban areas, and accurate DSMs are not always available. The U-Net model proposed in [27] is an improved FCN model, characterized by symmetrical U-shaped architecture consisting of an encoder and a decoder. This model combines low-level features with detailed spatial information with high-level features with semantic information to improve segmentation accuracy and achieves promising results in segmentation problems of individual classes, such as biomedical image segmentation, aerial road network definition, building identification and sea and land segmentation in Google Earth images. For multi-class classification tasks such as land cover classification, contextual information at different scales is important, as the characteristics of different types of land cover or terrestrial objects usually have different scales [27]. However, this information is not included in the original form of the U-Net model. Convolution is an important step in the CNN and FCN models because it allows models to distinguish features of different scales and degrees of abstraction.

2.3 NN Architecture

The initial form of the U-Net architecture consists of a compression path and a mirror expansion path. The convolution path determines high-level features through convolution and merges operations, while the spatial resolution of object maps is reduced. The expansion path (decoder) attempts to restore the resolution of object maps using up convolution operations. For each level of the compression path, feature maps are transmitted to a similar level in the decoder, which allows you to distribute contextual information over the network. For our experiments, we use one of the most accurate deep learning architectures for semantic segmentation tasks, the U-Net model. The architecture of our U-Net model traditionally consists of convolutional and de-convolutional parts, which are interconnected by a concatenation operation.

Traditionally, cross-entropy (CE) or weighted cross-entropy is chosen as a loss function (1) for multilayer perceptron training and deep learning models [5]:

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

where C is the number of classes, N is the number of elements in the sample, y is the target vector, in our case we choose one-hot coding, p is the output of the last layer of the neural network, α_c is the coefficient to control class influence, $\alpha_c = \frac{N_c}{N}$.

Batch normalization is applied before each Relu activation function. Batch normalization scales the input for layers, typically mini-packets, using mean value and variance. This scaling eliminates the internal covariance shift and thus speeds up the learning process [28]. A 2 × 2 max merge operation is used between the top and bottom layers on the compression path, which reduces the resolution of object maps. The size of the objects in the bottom layer of the compression contour is reduced to 1/64 of the original image. The bottom layer of the standard U-Net model corresponds to the same structure of the bottom layer, that is it includes two consecutive convolutions, batch normalization, and Relu operation.

2.4 Filtration Method

Given the high segmentation of the terrestrial cover map obtained as a result of the classification of high-resolution satellite data, it is necessary to improve the quality of the classification map by post-processing, namely the filtering of the resulting map. To do this, it is proposed to use an additional channel, which contains the recognition probabilities of each pixel according to each type of land cover.

The neural network used to construct the terrestrial cover map uses the Softmax layer to convert objects to the probability that a pixel belongs to each type of land cover. This function reduces the K-dimensional vector z with arbitrary values of the components to the K-dimensional vector σ (z) with the actual values of the components on the interval [0, 1] giving a sum of one. The function is set as follows (2), (3):

$$\sigma: \mathbb{R}^K \to [0,1]^K\,, \qquad (2)$$

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}}\,. \qquad (3)$$

The probability that an object x_i with a class y_i will occur in the sample is (4):

$$b(x_i)^{[y_i=+1]}(1 - b(x_i)^{[y_i=-1]}). \qquad (4)$$

Therefore, we will record the plausibility of the sample (the probability of obtaining such a sample in terms of algorithm) (5):

$$Q(a,X) = \prod_{i=1}^{l} b(x_i)^{[y_i=+1]}(1 - b(x_i)^{[y_i=-1]}), \qquad (5)$$

where X is the space of objects x, y_i = + 1 is the identifier that object x belongs to class +1.

This plausibility can be used as a functional for learning the algorithm, with an amendment that is more convenient to optimize its logarithm (6):

$$-\sum_{i=1}^{l}([y_i = +1]\log b(x_i) + [y_i = -1] \log(1 - b(x_i))) \to min. \qquad (6)$$

This loss function allows you correctly to predict the probabilities. Due to the need to filter the resulting mask of artificial objects from unreliable areas that erroneously fell into this mask due to classification, the method of threshold filtering (according to the principle of threshold classification [29]) was used with a threshold of 80% probability channel for artificial class objects.

2.5 Accuracy Assessment

Due to the limited resolution of satellite images, the density of buildings within quarters on Sentinel classification maps will be higher than they really are. That is why within the research framework, Planet satellite data with a spatial resolution of 3 meters are used to build a land cover classification map for the city of Kyiv. However, even in the absence of such data, for the construction of an urban atlas, the accuracy of classification according to Sentinel images is more than 90% [29], [30].

The condition for the creation of the Urban Atlas is the need to determine the percentage of buildings in each of the city's neighborhoods. This uses a map of the land cover and a vector layer with quarters of the city, built on the principle described above. To determine the percentage of the area of a certain type of land use (n - land cover class number) for each quarter k we will use the (7):

$$\forall k = \overline{1, QN}, n = \overline{1, CN}: \; P_k^n = \frac{S_k^n}{S_k} \cdot 100\%, \quad (7)$$

where QN - the number of quarters; CN - the number of earth cover classes on the map obtained as a result of neural network training; P^k_n - the ratio of the area of class n to the area of quarter k; S^n_k - the area of class n in quarter k; S_k - the area of quarter k. To determine the percentage of buildings, we use formula (7) for the class of land cover "artificial objects" in each of the quarters of the city of Kiev. According to the above formula, the percentages of building density by quarters of the city of Kyiv are calculated.

2.6 Results and Discussion

To build a map of the land cover for the Kyiv city, cloudless satellite images of Planet with a spatial resolution of 3 meters and 4 multispectral channels (Blue, Red, Green, and Near-Infrared), obtained on September 6, 2020, and June 11, 2020 (Figure 2), over urban areas of Kyiv were used. Python programming language with a set of libraries that work with geospatial objects, including GeoPandas, was used to analyze spectral channels and highlight geospatial features. For each of the types of land cover, based on the training data, raster images are created for each of the image channels of the Planet satellite.

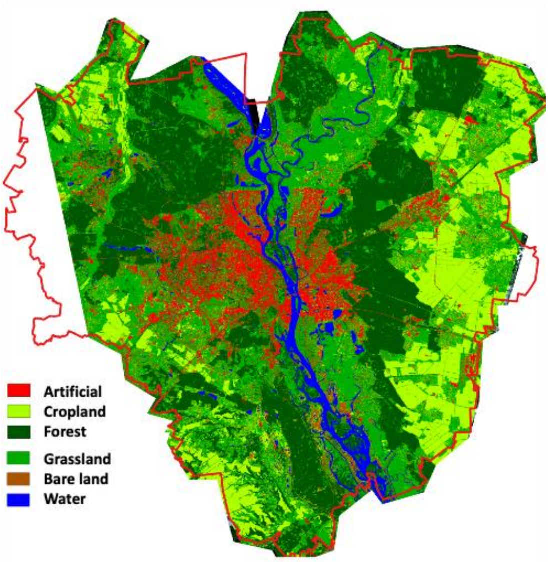

Figure 2: Classification map for the city of Kyiv and its environs according to Planet (2020).

Using the proposed neural network architecture [31], a map of the land cover classification for Kyiv and its environs was built according to Planet data with a spatial resolution of 3 meters (Figure 2). The overall accuracy of the obtained map is 95%, but there are some disadvantages due to the high spatial resolution of the data used.

3 CONCLUSIONS

The article proposes models for the classification of urban land cover by images with high spatial resolution. Like other supervised learning classification methods, the classification procedure consists of a learning phase and a testing phase. Unlike conventional learning strategies, which use randomly selected pixels or image objects as a training sample, the training patterns in our model are pairs of image fragments and their corresponding base chains, with each pixel marked with a specific class. Optimal sets of parameters are studied using inverse propagation and iterations [36]. In the classification phase, the trained models are executed on the input image to predict the class for each pixel.

The improved method of classification of land cover types on the basis of satellite data of high spatial resolution is described. The design process of information features for land cover types, the architecture of convolutional neural network for classification of satellite data of high spatial resolution is developed. The method of filtering artificial objects and other types of land cover using the probabilistic threshold method is proposed, and the problem of creating building density maps for city atlas quarters is solved.

Information features based on satellite data of high spatial resolution for different land cover types for land cover classification of the city of Kyiv are singled out and investigated. The architectures of neural networks used in the world for similar classification tasks are studied, the architecture of the neural network for segmentation and classification of satellite data of high spatial resolution for the city of Kyiv is developed. Using additional probabilistic information on the recognition of land cover classes, a method of filtering the obtained land cover classification map was developed. A method for estimating the density of the buildings within urban neighborhoods based on the obtained map of land cover classification has been developed. Metrics for the accuracy estimation of the obtained results are determined. All these methods and metrics are implemented in the Python programming language.

ACKNOWLEDGMENTS

This research was funded by the National Research foundation of Ukraine within the project 2020.02/0284 «Geospatial models and information technologies of satellite monitoring of smart city problems», which won the competition "Leading and Young Scientists Research Support".

REFERENCES

[1] P. Casals-Carrasco, S. Kubo, and B. Babu, "Madhavan Application of Spectral Mixture Analysis for Terrain Evaluation Studies", Int. J. of Remote Sensing, 2000, vol. 21, pp. 3039-3055.

[2] R. Manandhar, I. Odeh, and T. Ancev, "Improving the accuracy of land use and land cover classification of landsat data using post-classification enhancement", Remote Sens, 2009, vol. 1, pp. 330-344.

[3] T. Blaschke, "Object based image analysis for remote sensing", ISPRS J. Photogramm, Remote Sens, 2010, vol. 65, pp. 2-16.

[4] D. C. Duro, S. E. Franklin, and M. G. Dubé, "A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery", Remote Sens, Environ, 2012, vol. 118, pp. 259-272.

[5] Z. Y. Lv, P. L. Zhang, and J. A. Benediktsson, "Automatic object-oriented, spectral-spatial feature extraction driven by Tobler's first law of geography for very high-resolution aerial imagery classification", Remote Sens, 2017, vol. 9, p. 285.

[6] D. Marmanis, K. Schindler, J. D. Wegner, S. Galliani, M. Datcu, and U. Stilla, "Classification with an edge: Improving semantic image segmentation with boundary detection", ISPRS J. Photogramm, Remote Sens, 2018, vol. 135, pp. 158-172.

[7] W. Dong, T.J. Wu, J.C. Luo, Y.W. Sun, and L. G. Xia, "Land-parcel-based Digital Soil Mapping of Soil Nutrient Properties in an Alluvial-diluvia Plain Agricultural Area in China", Geoderma, 2019, vol. 340, pp. 234-248.

[8] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", In Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015), San Diego, CA, USA, 7-9 May 2015, pp. 1-14.

[9] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance", Int. J. Remote Sens, 2007, vol. 28, pp. 823-870.

[10] M. Pesaresi, A. Gerhardinger, and F. Kayitakire, "A robust built-up area presence index by anisotropic rotation-invariant textural measure", IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens, 2008, vol. 1, pp.180-192.

[11] D. J. Marceau, P. J. Howarth, J. M. Dubois, and D. J. Gratton, "Evaluation of the grey-level co-occurrence matrix method for land-cover classification using SPOT imagery", IEEE Trans. Geosci, Remote Sens, 1990, vol. 28. pp. 513-519.

[12] E. Hussain and J. Shan, "Object-based urban land cover classification using rule inheritance over very high-resolution multisensor and multitemporal data", GISci, Remote Sens, 2016, vol. 53, pp. 164-182.

[13] D. Li, Y. Ke, H. Gong, and X. Li, "Object-based urban tree species classification using bi-temporal WorldView-2 and WorldView-3 images", Remote Sens, 2015, vol. 7. pp. 16917-16937.

[14] G. Fu, H. Zhao, C. Li, and L. Shi, "Segmentation for High-Resolution Optical Remote Sensing Imagery Using Improved Quadtree and Region Adjacency Graph Technique", Remote Sens, 2013, vol. 5, pp. 3259-3279.

[15] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, W. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding", In Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016, pp. 3213-3223.

[16] V. Mnih and G. E. Hinton, "Learning to detect roads in high-resolution aerial images", In Computer Vision—ECCV 2010, Lecture Notes in Computer Science; Daniilidis, K., Maragos, P., Paragios, N., Eds., Springer: Berlin/Heidelberg, Germany, 2010, vol. 6316, pp. 210-223.

[17] J. Wang, J. Song, M. Chen, and Z. Yang, "Road network extraction: A neural-dynamic framework based on deep learning and a finite state machine", Int. J, Remote Sens, 2015, vol. 36, pp. 3144-3169.

[18] F. Hu, G. S. Xia, J. Hu, and L. Zhang, "Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery," Remote Sens, 2015, vol. 7, pp. 14680-14707.

[19] M. Längkvist, A. Kiselev, M. Alirezaie, and A. Loutfi, "Classification and segmentation of satellite orthoimagery using convolutional neural networks", Remote Sens, 2016, vol. 8, p. 329.

[20] E. Maltezos, "Deep convolutional neural networks for building extraction from orthoimages and dense image matching point clouds", J. Appl, Remote Sens, 2017, vol. 11, pp. 1-22.

[21] X. Pan and J. Zhao, "A central-point-enhanced convolutional neural network for high-resolution remote-sensing image classification", Int. J, Remote Sens, 2017, vol. 38, pp. 6554-6581.

[22] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 5-7 June 2015, pp. 3431-3440.

[23] J. Sherrah, "Fully convolutional networks for dense semantic labeling of high-resolution aerial imagery", arXiv, 2016, arXiv:1606.02585.

[24] N. Audebert, B. Le Saux, and S. Lefè vre, "Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks", ISPRS J. Photogramm, Remote Sens, 2018, vol. 140, pp. 20-32.

[25] X. Sun, S. Shen, X. Lin, and Z. Hu, "Semantic Labeling of High Resolution Aerial Images Using an Ensemble of Fully Convolutional Networks", J. Appl, Remote Sens, 2017, vol. 11, p. 042617.

[26] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "High-resolution aerial image labeling with convolutional neural networks", IEEE Trans. Geosci, Remote Sens, 2017, vol. 55, pp. 7092-7103.

[27] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation", In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5-9 October 2015, Springer: Cham, Switzerland, 2015, pp. 234-241.

[28] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift", arXiv, 2015, arXiv:1502.03167.

[29] M. Lavreniuk, N. Kussul, and A. Novikov "Deep learning crop classification approach based on coding input satellite data into the unified hyperspace", 38th International Conference on Electronics and Nanotechnology (ELNANO), 2018, pp. 239-244, doi: 10.1109/ELNANO.2018.8477525.

[30] M. Lavreniuk, N. Kussul, and A. Novikov "Deep Learning Crop Classification Approach Based on Sparse Coding of Time Series of Satellite Data", In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, 2018, pp. 4812-4815, doi: 10.1109/IGARSS.2018.8518263.

[31] A. Shelestov, H. Yailymova, B. Yailymov, L. Shumilo, and A. Lavreniuk "Extension of Copernicus Urban Atlas to non-european countries", 2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2021, Brussels (virtual format), pp. 6789-6792, doi: 10.1109/IGARSS47720.2021.9553546.

[32] G. M. Bakan and N. N. Kussul, "Fuzzy ellipsoidal filtering algorithm of static object state", Problemy Upravleniya I Informatiki (Avtomatika), 1996, no. 5, pp. 77-92.

[33] A. N. Kravchenko, N. N. Kussul, E. A. Lupian, V. P. Savorsky, L. Hluchy, and A. Y. Shelestov, "Water resource quality monitoring using heterogeneous data and high-performance computations", Cybernetics and Systems Analysis, 2008, vol. 44(4), 616-624. doi:10.1007/s10559-008-9032-x.

[34] N. Kussul, A. Shelestov, B. Yailymov, H. Yailymova, M. Lavreniuk, L. Shumilo, and Y. Bilokonska, "Crop monitoring technology based on time series of satellite imagery", 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies, DESSERT 2020, May 2020, pp. 346-350.

[35] N. Kussul, A. Shelestov, H. Yailymova, B. Yailymov, M. Lavreniuk, M. Ilyashenko, "Satellite Agricultural Monitoring in Ukraine at Country Level: World Bank Project", 2020 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2020, 26 September 2020, pp. 1050-1053.

[36] A. Shelestov, A. Kolotii, S. Skakun, B. Baruth, R.L. Lozano, and B. Yailymov, "Biophysical parameters mapping within the SPOT-5 take 5 initiative", European Journal of Remote Sensing, vol. 50, Issue 1, 2017, pp. 300-309.

[37] N. Kussul, A. Kolotii, A. Shelestov, B. Yailymov, and M. Lavreniuk, "Land degradation estimation from global and national satellite based datasets within un program", 2017 IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2017, vol. 1, 3 November 2017, pp. 383-386.