DEEP LEARNING APPROACH FOR CROP MAPPING ON BIG DATA

Nataliia Kussul
National Technical University of Ukraine Kyiv Polytechnic Institute

Andrey Yu. Shelestov
National Technical University of Ukraine Kyiv Polytechnic Institute

Mykola Lavreniuk
National Academy of Sciences of Ukraine

Bohdan Yailymov
Space Research Institute NAS Ukraine and SSA Ukraine

NATIONAL ACADEMY OF SCIENCES OF UKRAINE

SPACE RESEARCH IN UKRAINE

2016—2018

Report
to COSPAR

The Report Prepared by the Space Research Institute of NAS of Ukraine and SSA of Ukraine

Scientific Editor O. FEDOROV

KYIV • AKADEMPERIODYKA • 2018

UDK 001.891(15)
BBK 22.6
S78

SRI NASU-SSAU
Space Research Institute
of National Academy of Sciences of Ukraine
and State Space Agency of Ukraine
Ukraine
03187, Kyiv
40, bd 4/1, Acad. Glushkov av.
http://www.ikd.kiev.ua

NASU
National Academy of Sciences of Ukraine
Ukraine
01601, Kyiv 30
54, Volodymyrska St.
http://www.nas.gov.ua

Space research in Ukraine. 2016–2018 / Ed. O. Fedorov. — K.:
S78 Akademperiodyka, 2018. — 164 p.
ISBN 978-966-02-8590-3

Ukrainian report to COSPAR summarizes the results of space research performed during the years 2016—2018. This edition presents the current state of Ukrainian space science in the following areas: astronomy, Earth exploration and near-Earth space research, life sciences, space technologies and materials science. A number of papers are dedicated to the creation of scientific instruments for perspective space missions. Considerable attention paid to applied research of space monitoring of the Earth. The collection is the Ukrainian report to COSPAR.

The collection can be useful for a wide range of readers, interested in space research.

UDK 001.891(15)
BBK 22.6

ISBN 978-966-02-8590-3

DEEP LEARNING APPROACH FOR CROP MAPPING ON BIG DATA

N. Kussul[1,2], A. Shelestov[2,1], M. Lavreniuk[1,2], B. Yailymov[1], A. Kolotii[1,2], H. Yailymova[1,3], S. Skakun[4], L. Shumilo[1,2]

[1] Space Research Institute of NAS of Ukraine and SSA of Ukraine (SRI NASU-SSAU)

[2] National Technical University of Ukraine «Kyiv Polytechnic Institute» (NTUU «KPI»)

[3] Taras Shevchenko National University of Kyiv (KNU)

[4] University of Maryland, College Park, MD, USA

The last several years and onwards could be called the years of Big Free Data in domain of Remote Sensing. During the 2013–2016 period, several optical and synthetic aperture radar (SAR) remote sensing satellites were launched with high spatial resolution (10–30 m), in particular Sentinel-1A/B and Sentinel-2A within the European Copernicus program [1]. These datasets are freely available on operational basis. This opens unprecedented opportunities for a wide range of pre-operational and operational applications in the environment and agricultural domains taking advantage of high temporal and spatial resolution datasets and advances in the multi-sources data fusion techniques [2–4]. Due to clouds and shadows, the amount of available optical data over the region of interest is limited. SAR Sentinel-1 (both A and B) data are weather independent and new opportunities for crop classification have opened.

Crop mapping based on high resolution satellite data is a very important component for solving various applied problems, in particular crop area estimation [4], yield forecasting [5] and drought risk quantification [6–9]. Earlier, SAR data were quite expensive, infrequent and most of the studies on crop state assessment and crop type mapping were performed with optical data only. High spatial (10 m) and temporal (6 days revisit) resolution of the Sentinel-1 mission brings new opportunities in the agriculture domain and challenges of «Big data» problems in Remote Sensing that should be addressed.

In this paper, we propose a multi-layer Deep Learning (DL) architecture that is targeted for classification of multi-source multi-temporal remote sensing images, both optical and SAR, at a pixel level [10, 11, 12]. The core of the architecture is an ensemble of convolutional neural networks (CNNs) [13, 14]. The proposed architecture is applied for crop classification using Sentinel-1A time-series and provides accuracy high enough to be considered for operational context at the national level.

Data description

We address the problem of land cover and crop classification for Kyiv region of Ukraine using multi-temporal multi-source images acquired by Sentinel-2 and Sentinel-1A satellites [15]. The study area is classified into eleven classes including major agricultural crops (water, forest, grassland, bare land, winter wheat, winter rapeseed, spring cereals, soybeans, maize, sunflowers and sugar beet). It is rather large area (28.000 square km) with big diversity of different land cover types and agricultural crops. The territory is big enough to be considered as a representative one for the extension of the technology to the entire country. For the 2016 vegetation season (since October 2015 till September 2016) five cloud-free Sentinel-2 and twenty Sentinel-1 images were used for the study area. Sentinel-1A images were preprocessed via following steps: calibration, multi-looking (with 2×2 window), speckle filtering (3×3 window with Refined Lee algorithm), and terrain correction using SRTM DEM. A time-series of four 10 meter spectral bands from each Sentinel-2 scene and two bands with VV and VH polarizations from each Sentinel-1 scene are used as an input to the classification model.

For crop classification for the territory of Ukraine time series of only SAR Sentinel-1 data [16, 17, 18] were used. For the territory of Ukraine 9 paths of Sentinel-1 were used. Each of these paths consist of several images (Table 1, Fig. 1) and is constructed by merging them for one date in single stripe.

Table 1

The data size needed to crop mapping for the territory of Ukraine

Ukraine			
Path	Number of images for each path	Number of path 2016	Size of one Path, GB
7	4	19	7.19
21	3	21	2.32
36	6	20	11.31
65	6	19	12.08
80	4	19	7.14
94	5	20	8.26
109	3	22	6.69
138	5	22	11.89
167	5	23	10.67
Total	41	185	1576.08

More than 800 images were used for covering the territory of Ukraine with Sentinel-1A data during the vegetation season. The data amount used for land cover classification for 2016 is over 1.5 Tb in total.

Crop type and cropland mapping datasets were collected during 2-stage along the road surveys separately for winter crops (in the spring of 2016) and summer crops (in the summer of 2016) – 5526 samples for calibration and 2154 samples for validation purposes for major cropland and non-cropland classes – Fig. 2.

Non-cropland classes were added using both in-situ surveys and photo-interpretation from high-resolution imagery from Google Earth. Data collection was performed according to climatic zonation of Ukraine and with respect to JECAM guidelines – within 4 main climatic stratas of Ukraine (Woodlands, Forest and Steppe, Steppe and Mountains) – Fig. 3.

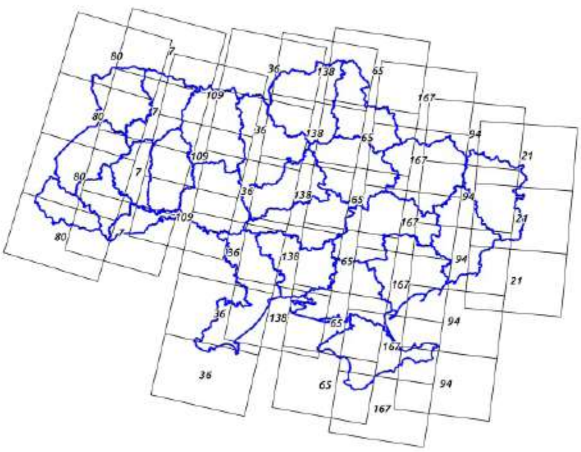

Fig. 1. Coverage of the territory of Ukraine by Sentinel-1A data with relative orbit number

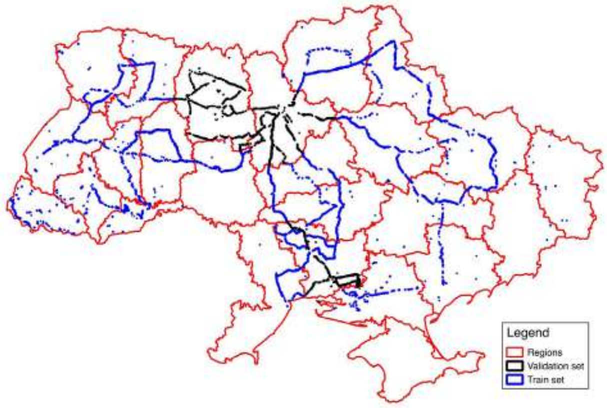

Fig. 2. In-situ data distribution over the territory of Ukraine

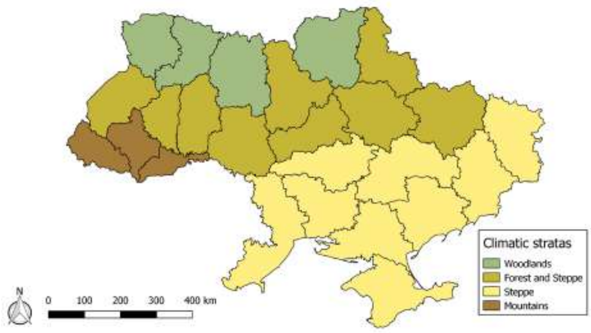

Fig. 3. Main climatic stratas of Ukraine

Methodology and products

When providing large scale crop mapping using multi-temporal satellite imagery, the following challenges should be addressed while using DL. First, pixels of a satellite image contain physical values. In particular, each pixel of the optical imagery contains spectral reflectance values in multiple spectral bands, and can be contaminated with clouds and shadows; while each pixel of the space-borne SAR imagery is characterized by backscatter intensity and phase in multiple polarizations. Both of the data sources have multi-temporal nature and different spatial resolutions. That is why, DL implementation for land cover and crops classification based on data fusion of multi-temporal multi-sensor satellite data is a challenge.

A four-level architecture is proposed for classification of crop types from multi-temporal satellite imagery. These levels are pre-processing, supervised classification, post-processing and geospatial analysis (Fig. 4).

Since optical satellite imagery can be contaminated with clouds and shadows, one have to deal with missing values in the imagery. Most classifiers accept only valid pixel values as an input, and therefore a pre-processing step should be performed to

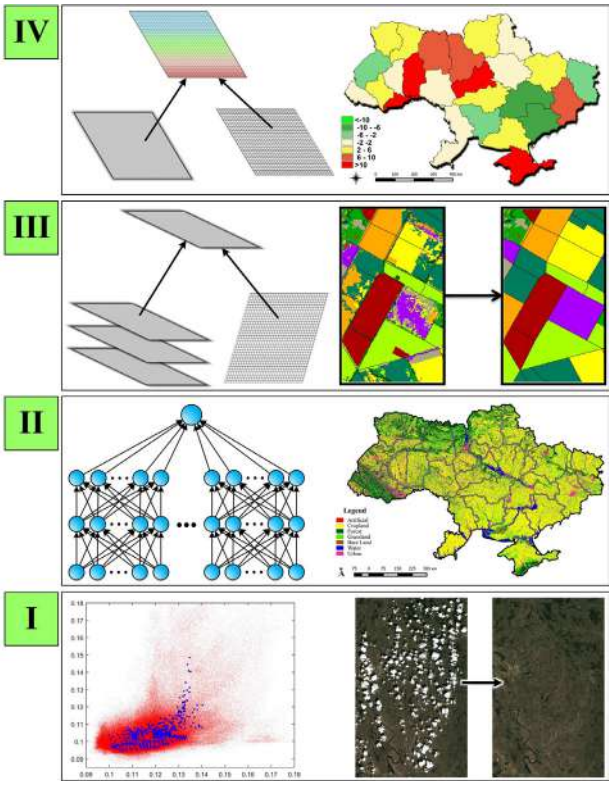

Fig. 4. A four-level hierarchical deep learning model for satellite data classification and land cover/land use changes analysis (I – pre-processing for dealing with missing data on optical images due to clouds /shadows, II – supervised classification, III – post-processing using additional geospatial data to improve classification maps, IV – geospatial analysis for a high-level product, e.g. crop area estimation)

impute (or fill gaps) missing values. This procedure is performed within level I of the architecture. The next step is supervised classification (level II) which is the core of this study. We propose CNNs architectures, namely 2-d, to explore spectral and spatial features, respectively. Levels III and IV are aimed at improving the resulting classification map with available geospatial layers and building high-level products.

For pre-processing, we utilize self-organizing Kohonen maps (SOMs) for optical images segmentation and subsequent restoration of missing data in a time-series of satellite imagery [15]. SOMs are trained for each spectral band separately using non-missing values. Missing values are restored through a special procedure that substitutes input sample's missing components with neuron's weight coefficients. Pixels that have been restored are masked, the number of cloud-free scenes available for each pixel from optical imagery is calculated, and these two layers are used for further post-processing procedure (at level III) to improve the resulting classification map [19]. The detailed description of the restoration algorithm is given in [2, 15].

The core element of the model is the supervised classification, which is performed at the second stage (level II). The CNN, in turn, builds a hierarchal set of features through local convolution and down-sampling. The two bands from each of the twenty Sentinel-1A scenes and the four bands from each of the four Sentinel-2 scenes form a CNN input feature vector with dimension size 60 ($20 \times 2 + 5 \times 4$). Traditional CNNs (2-d) take into account a spatial context of an image and provide higher accuracy comparing to a per pixel-based approach. However in this case, CNN smooths not only some misclassified pixels but also small objects like roads, and forest «stripes» and clear cuts within the forest (with linear dimensions of several pixels) are missed. Each CNN in the corresponding ensemble consists of two convolutional layers, each of them followed by max-pooling and two fully connected layers in the end (Fig. 5). We used a rectified linear unit (ReLU) function that is one of the most popular and efficient activation functions for deep neural networks. There are advantages of using ReLU such as biological plausibility, efficient computation and gradient propagation. Therefore, ReLU function is faster and more effective for training CNNs comparing to a sigmoid function. Each of the CNNs has the same convolution and max-pooling structure but differs in the trained filters and number of neurons in the hidden layer (NNH) being 60, 70, 80, 90 and 100 for five CNNs, respectively.

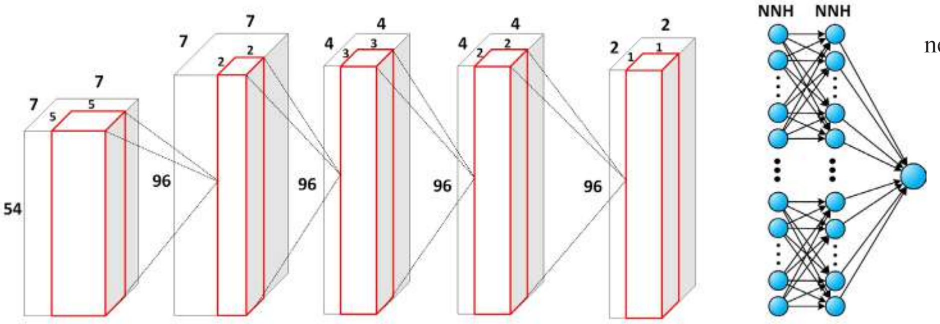

Fig. 5. Deep convolutional neural network architecture

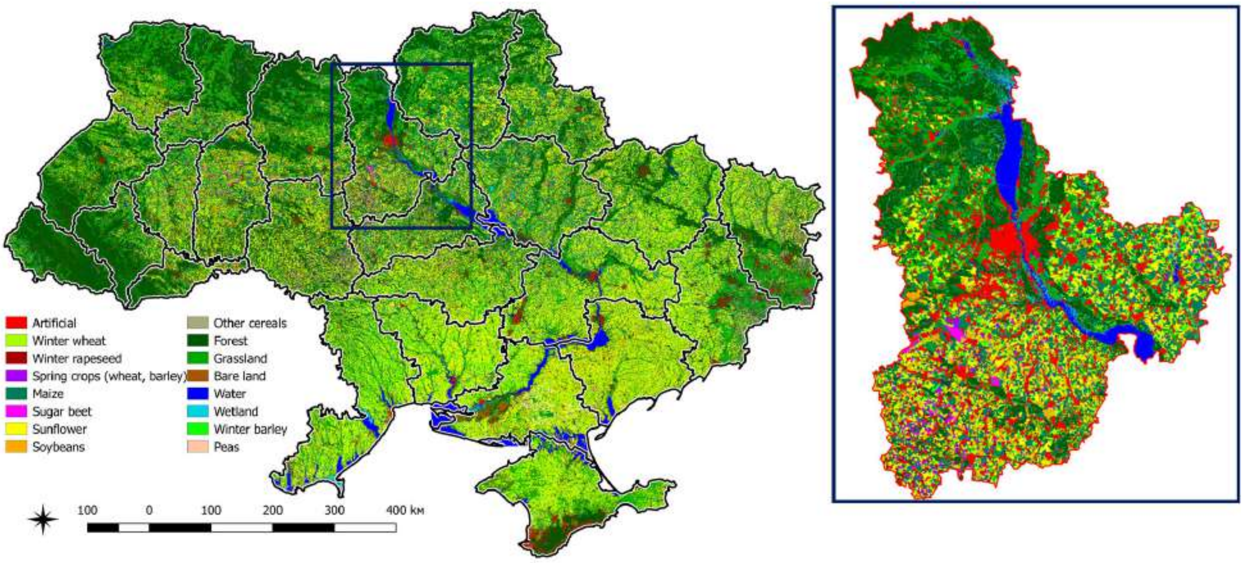

Fig. 6. Crop type map for Ukraine,2016 (SRI product)

To improve the quality of the resulting map, we developed several filtering algorithms, based on the available information on quality of input data and fields boundaries [19]. Those filters take a pixel-based classification map and specifically designed rules to account for several plots (fields) within the parcel. In the result, we obtained a clear parcel-based classification map. The final level of data processing provides data fusion with multi sourced heterogeneous information, in particular, statistical data, vector geospatial data, socio-economic information and so on. It allows interpreting the classification results, solving applied problems for different domains, and providing the support information for decision makers. For example, classification map coupled with area frame sampling approach can be used to estimate crop areas [6].

Results analysis

With use of the proposed four-level architecture were in Space Research Institute (SRI) obtained crop classification map for Kiev region and for territory of Ukraine. These maps were validated on independent in-situ data from ground survey. The overall accuracy of classification map for Kiev region is over 93% and over 90% for Ukraine in 2016. Land cover maps in the end of vegetation period for territory of Ukraine is shown on Fig. 6. For main crops the F1-score has been calculated based on independent test set: winter wheat – 90,

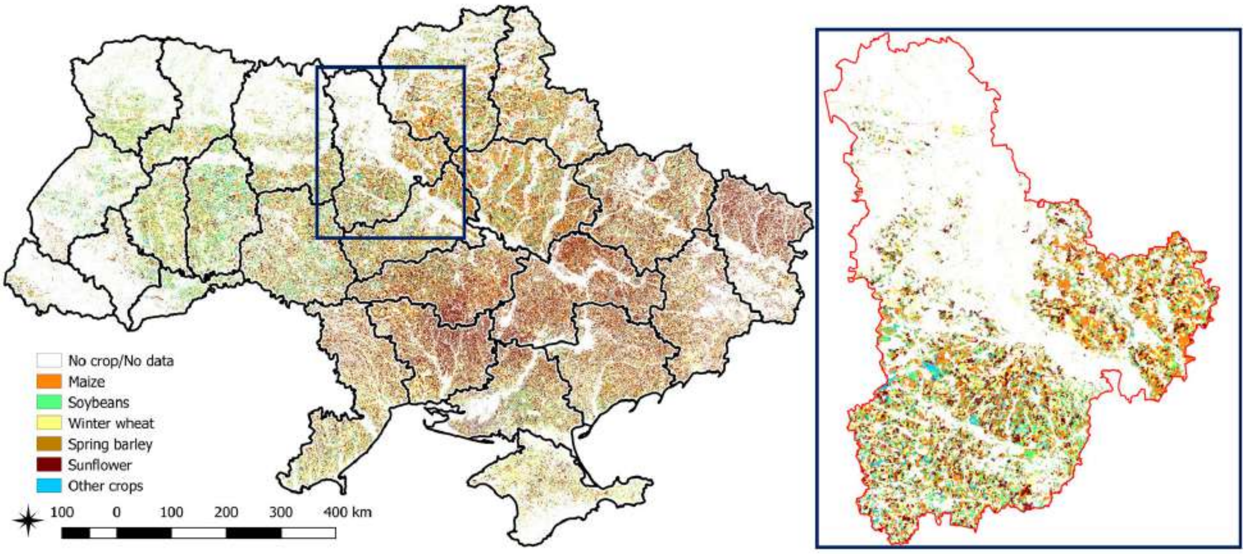

Fig. 7. Crop type map for Ukraine, 2016 (Sen2-Agri product)

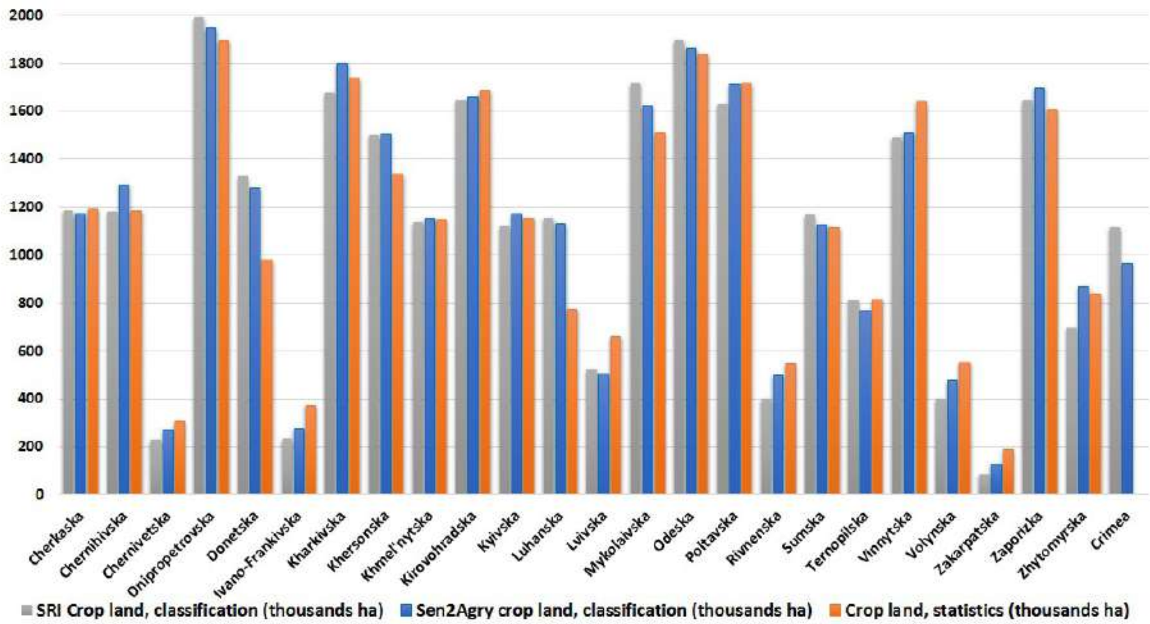

Fig. 8. Crop area comparison to official statistics

winter rapeseed – 83.6, maize – 93, sugar beet – 93.6, sunflower – 94.3, soybeans – 82.5, peas – 70.9.

In this section SRI crop type map for Ukraine will be compared to statistical information from State Statistical Service of Ukraine. Cross-comparison of SRI crop type map developed with use of deep learning and alternative Sen2-Agri crop type map (Fig. 7) will be performed as well (Fig. 8).

According to official data the biggest difference in cropland area correspond to the conflict area in the Eastern part of Ukraine (Donetska and Luhanska regions – Fig. 8) and caused by several reasons: (1) non-fully coverage of the area be statistical observations; (2) significant decrease of agricultural activity over this areas [13].

Discrepancies in the Western part of Ukraine is common for this region and caused by the specifics of agriculture and statistical data collection for these regions.

Table 2

Cross-comparison of crop type maps with official statistics over pilot areas for major crops

Crop type	Areas in thousands of hectares		
	SRI	Sen2-Agri	Statistics
	Kiev region		
Maize	314.9	312.6	268
Winter wheat	204.5	262.4	197.2
	Mykolaiv region		
Maize	99.7	166.4	132.8
Winter wheat	480.6	481.0	331.8

For main crops (maize and winter wheat) for two regions we performed cross-comparison of SRI and Sen2-Agri crop type maps with official statistics – Table 2. In Kiev region results are highly agreed for maize, in Mykolaiv region – for winter wheat.

Conclusions

In this paper, we proposed a multi-level Deep Learning approach for land cover and crop types classification using multi-temporal multi-source satellite imagery. The architecture uses both unsupervised and supervised neural networks for segmentation and subsequent classification of satellite imagery, respectively. In this study, we used Sentinel-2 and Sentinel-1A images for Kiev region and all Ukraine. In general, the use of CNN allowed us to reach the target accuracy over 90 % for all territory of Ukraine in 2016 (by 9 % more than the results of Sen2-Agri national demonstration). In the future it is planned to implement the developed architecture in the cloud [20] for the operational updating of land cover maps for Ukraine.

REFERENCES

1. *Kussul N., Kolotii A., Shelestov A., Lavreniuk M., Bellemans N., Bontemps S., Defourny P., Koetz B.* Sentinel-2 for agriculture national demonstration in Ukraine: Results and further steps // IEEE International Geoscience and Remote Sensing Symposium (IGARSS). – 2017. – P. 5842–5845. DOI: 10.1109/IGARSS.2017.8128337.

2. *Kussul N., Skakun S., Shelestov A., Lavreniuk M., Yailymov B., Kussul O.* Regional scale crop mapping using multi-temporal satellite imagery // International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences. – 2015. – P. 45–52.

3. *Lavreniuk M.S., Skakun S.V., Shelestov A.Ju., Yalimov B.Ya., Yanchevskii S.L., Yaschuk D.Ju. and Kosteckiy A.M.* Large-Scale Classification of Land Cover Using Retrospective Satellite Data // Cybernetics and Systems Analysis. – 2016. – V. 52. – No. 1. – P. 127–138. DOI: 10.1007/s10559-016-9807-4.

4. *Francois W., Kussul N., Guerric le Maire, Dupuy S.* Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity // International Journal of Remote Sensing. – July 2016. – V. 3714, No. 14. – P. 3196–3231. DOI: 10.1080/01431161.2016.1194545.

5. *Kogan F., Kussul N., Adamenko T., Skakun S., Kravchenko O., Kryvobok O., Shelestov A., Kolotii A., Kussul O., Lavrenyuk A.* Winter wheat yield forecasting: A comparative analysis of results of regression and biophysical models // Journal of Automation and Information Sciences. – 2013. – V. 45. – No. 6. – P. 68–81.

6. *Gallego J., Kussul N., Skakun S., Kravchenko O., Shelestov A., Kussul O.* Efficiency assessment of using satellite data for crop area estimation in Ukraine // International Journal of Applied Earth Observation and Geoinformation. – 2014. – V. 29. – P. 22–30.

7. *Kussul N., Skakun S., Shelestov A., Kussul O., Yailymov B.* Resilience Aspects in the Sensor Web Infrastructure for Natural Disaster Monitoring and Risk Assessment Based on Earth Observation Data // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. – 2014. – V. 7. – No. 9. – P. 3826–3832.

8. *Kussul N., Shelestov A., Skakun S.* Flood Monitoring from SAR Data // NATO Science for Peace and Security Series C: Environmental Security. – 2011. – P. 19–29. DOI: 10.1007/978-90-481-9618-0_3.

9. *Kravchenko A., Kussul N., Lupian E., Savorsky vV., Hluchy L., Shelestov A.* Water resource quality monitoring using heterogeneous data and high-performance computations // Cybernetics and Systems Analysis. – 2008. – V. 44. – No. 4. – P. 616–624. DOI:10.1007/s10559-008-9032-x.

10. *Kussul N., Shelestov A., Skakun S., Kravchenko O.* High-performance intelligent computations for environmental and disaster monitoring // Int. J. Information Technologies & Knowledge. – 2009. – V. 3. – P. 135–156.

11. *Kussul N., Shelestov A., Basarab R., Skakun S., Kussul O., Lavreniuk M.* Geospatial intelligence and data fusion techniques for sustainable development problems // 11th International Conference on ICT in Education, Research and Industrial Applications: Integration, Harmonization and Knowledge Transfer (ICTERI 2015), May 14–16, 2015, Lviv, Ukraine. – 2015. – V. 1356. – P. 196–203.

12. *Lavreniuk M., Kussul N., Skakun S., Shelestov A., Yailymov B.* Regional Retrospective High Resolution Land Cover For Ukraine: Methodology And Results // International Geoscience and Remote Sensing Symposium 2015 (IGARSS 2015), № 15599383. – P. 3965–3968. DOI: 10.1109/IGARSS.2015.7326693.

13. *Kussul N., Lavreniuk N., Shelestov A., Yailymov B., Butko I.* Land Cover Changes Analysis Based on Deep Machine Learning Technique // Journal of Automation and Information Sciences. – 2016. – V. 48. No. 5. – P. 42–54. DOI: 10.1615/JAutomatInfScien.v48.i5.40.

14. *Kussul N., Lavreniuk M., Skakun S., Shelestov A.* Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data // IEEE Geoscience and Remote Sensing Letters. – 2017. – V. 12. – No. 5. – P. 778–782. DOI: 10.1109/LGRS.2017.2681128.

15. *Skakun S.V., Basarab R.M.* Reconstruction of Missing Data in Time-Series of Optical Satellite Images Using Self-Organizing Kohonen Maps // Journal of Automation and Information Sciences. – 2014. – V. 46. – No. 12. – P. 19–26.

16. *Kussul N., Skakun S., Shelestov A., Kussul O.* The use of satellite SAR imagery to crop classification in Ukraine within JECAM project // IEEE International Geoscience and Remote Sensing Symposium (IGARSS). – 2014. – P. 1497–1500.

17. *Skakun S., Kussul N., Shelestov A.Y., Lavreniuk M. and Kussul O.* Efficiency Assessment of Multitemporal C-Band Radarsat-2 Intensity and Landsat-8 Surface Reflectance Satellite Imagery for Crop Classification in Ukraine // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. – 2016. – V. 9. – No. 8. – P. 3712-3719. DOI:10.1109/JSTARS.2015.2454297.

18. *Abramov S., Rubel O., Lukin V., Kozhemiakin R., Kussul N., Shelestov A., Lavreniuk M.* Speckle reducing for Sentinel-1 SAR data // IEEE International Geoscience and Remote Sensing Symposium (IGARSS). – 2017. – P. 2353–2356. DOI: 10.1109/IGARSS.2017.8127463.

19. *Kussul N., Lemoine G., Gallego J., Skakun S., Lavreniuk M.* Parcel based classification for agricultural mapping and monitoring using multi-temporal satellite image sequences // The International Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International. – 2015. – P. 165-168. DOI: 10.1109/IGARSS.2015.7325725.

20. *Kussul N., Shelestov A., Skakun S.* Grid technologies for satellite data processing and management within international disaster monitoring projects // Grid and Cloud Database Management. – 2011. – P. 279–305.

NATIONAL ACADEMY OF SCIENCES OF UKRAINE
STATE SPACE AGENCY OF UKRAINE

SPACE RESEARCH IN UKRAINE 2016—2018

The report was prepared by Space Research Institute of NAS of Ukraine and SSA of Ukraine

in English

Scientific editor
FEDOROV Oleg

Compiler
SAMOYLENKO Lyudmyla

Computer layout
T. Skorokhod