DEEP LEARNING APPROACH FOR LARGE SCALE LAND COVER MAPPING BASED ON REMOTE SENSING DATA FUSION

Nataliia Kussul[1, 2], *Andrii Shelestov*[1, 2], *Mykola Lavreniuk*[1], *Igor Butko*[1], *Sergii Skakun*[3]

[1]Space Research Institute NASU-SSAU, Kyiv, Ukraine
[2]National Technical University of Ukraine "Kyiv Polytechnic Institute", Kyiv, Ukraine
[3]University of Maryland, College Park, MD, USA

ABSTRACT

In the paper we propose the methodology for solving the large scale classification and area estimation problems in the remote sensing domain on the basis of deep learning paradigm. It is based on a hierarchical model that includes self-organizing maps (SOM) for data preprocessing and segmentation (clustering), ensemble of multi-layer perceptrons (MLP) for data classification and heterogeneous data fusion and geospatial analysis for post-processing. The proposed methodology is applied for generation of high resolution land cover and land use maps for the territory of Ukraine from 1990 to 2010 and 2015.

***Index Terms*—** Deep learning, neural network, remote sensing data, big data, geospatial analysis, Landsat

1. INTRODUCTION

An agreement under the United Nations (UN) Framework Convention on Climate Change signed during the COP21 meeting in Paris in December 2015[1] emphasizes on implementation of the Convention and recognizes the role of sustainable development as set in the UN Resolution on "Transforming our world: the 2030 Agenda for Sustainable Development"[2].

With the constantly increasing volume of remote sensing (RS) data, this becomes a powerful tool in addressing the challenges and improving our understanding of the Earth system. In particular, with launches of Sentinel-1, Sentinel-2, Proba-V and Landsat-8 satellites, there will be generated up to petabyte of raw images per year. These images and derived products are extremely important for many applications in climate change, environment, disaster monitoring and risk assessment, agriculture and food security. For example, the size of the Climate Change data repositories alone is projected to grow to nearly 350 Petabytes by 2030[3].

On the other hand, the increasing volume of remote sensing data, dubbed as a "Big Data" problem [1]–[3], creates new challenges in handling datasets that require new approaches to extracting relevant information from RS data from data science perspective [4]. Modern RS is no longer considered as a processing of a one source single date image — it has shifted to multi-source fusion of multi-temporal images. The data can be presented in the form of a 4D "data cube": 2D spatial component, multi- or hyper-spectral component, and temporal component. Moreover, each component might have its own "heterogeneity": discrepancy of spatial resolution of multi-source data (from high to low resolution); different spectral bands coming, for example, from Land Surface Imaging Constellation (LSI)[4]; and non-uniform temporal coverage. Therefore, the analysis should focus at different scales including different hierarchical levels and utilizing different semantics at various levels. An ensemble of methods ("mixture of experts" approach) should be exploited to take advantage of different processing methods and techniques. Advanced data processing techniques are to be developed and applied to address these critical challenges.

In past years there has been a large boost in developing advanced machine learning techniques, in particular deep learning [5]. Within this approach, we propose a hierarchical neural network based model for large scale classification of remote sensing images. The paper proposes to advance intelligent methods and technologies to fuse/integrate data and products acquired by multiple heterogeneous sources using machine learning techniques and emerging big data and geo-information technologies, to provide data processing and visualization capabilities, and their applications to social important areas. Possible applied areas include agriculture and food security, risk analysis, disaster monitoring and management, and environment monitoring.

[1] http://unfccc.int/resource/docs/2015/cop21/eng/l09.pdf
[2] https://sustainabledevelopment.un.org/post2015/transformingourworld/publication
[3] https://open.nasa.gov/blog/what-is-nasa-doing-with-big-data-today/
[4] http://ceos.org/ourwork/virtual-constellations/lsi/

2. METHODOLOGY

Deep learning techniques offer unique capabilities in extracting and presenting features at different levels from input data. These features are extremely important and should allow better presentation of remote sensing images in image classification/segmentation tasks.

In this paper we present the current results in exploring feasibility and assessing efficiency of applying deep learning techniques to classification and segmentation of multi-temporal multi-spectral and radar satellite images. For this purpose, we propose the following hierarchical model shown in Fig. 1. It includes four levels of data processing: preprocessing, classification, postprocessing and geospatial analysis for different applications.

1. Preprocessing (noise filtration and data clustering). Optical satellite data are often contaminated by clouds and shadows. The problem of their elimination is solved by clustering time-series of satellite images on the basis of self-organizing maps (SOM). In Fig. 1, *I* (left) satellite data, projected to the plane, are highlighted in red; results of clustering are highlighted in blue. An example of such kind of data processing for optical Landsat-8 data is shown in Fig. 1, *I* (right).

2. Land cover classification using time-series of satellite images. As it was shown in previous studies [9]–[11], one of the most efficient approaches to big data classification is an ensemble of neural networks. It can be used for classification of time-series of optical images as well as for optical and radar data integration.

3. Map postprocessing with filtering of salt and pepper and image vectorization. Since we use pixel-based classification, the resulting map may contain noise in the form of small groups of isolated misclassified pixels. To deal with this problem we propose to use combination of morphological and logical operations with median filtration [12].

4. Geospatial analysis of the results, received at previous levels. For this purpose we use auxiliary information, such as vector boundaries of some parcels, administrative regions, statistical data and so on [6, 13]. At this level we can solve sustainable development problems using regional statistical data and economical indicators of their development.

3. APPLICATIONS

The deep learning approach described in previous section can be applied to address social benefit areas such as land cover and land use (LCLU) change, disaster risk assessment, agriculture and food security [14].

With our approach the problem of land cover/land use (LCLU) and crop type classification is addressed using high-resolution (at 30 m spatial resolution) satellite imagery: Landsat-8, Sentinel-1 and Sentinel-2.

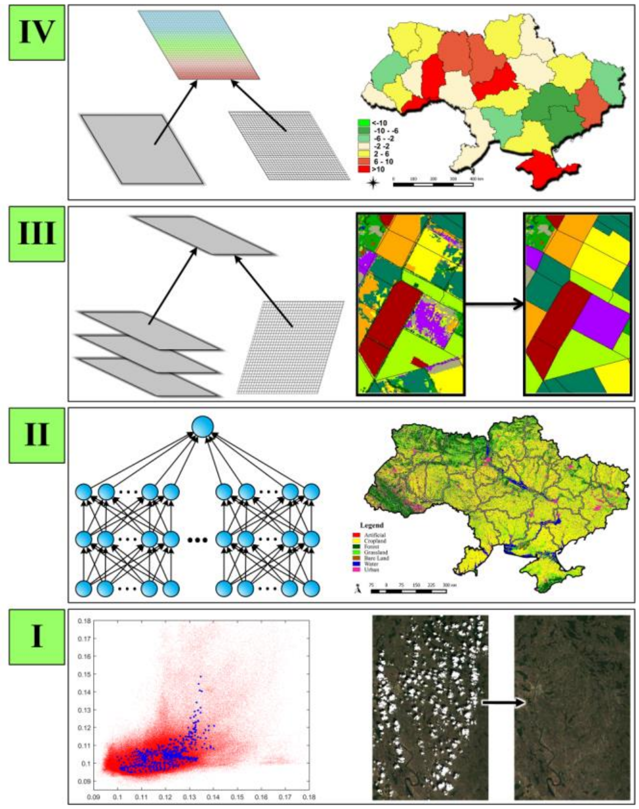

Fig. 1. A four-level hierarchical deep learning model for satellite data classification and land cover/land use changes analysis [6].

The proposed approach was applied to develop high resolution maps of land cover for Ukraine (1990, 2000, 2010 and 2015) and estimated land use changes for the last three decades. We found out that area of arable lands significantly decreased in Western and Northern parts of Ukraine and in the Eastern administrative regions (Fig 2). In the Fig. 2 it is shown the percentage of increasing non-cultivated agricultural lands (yellow to red colors) corresponding to the total area of region.

With global coverage and availability of Sentinel-1/SAR imagery, it becomes possible to develop innovative products by extracting useful information to serve social important applications. One such area includes agriculture and food security. The target is to improve capabilities relating to large scale crop mapping and crop area estimation; crop condition, yield and crop production forecasting; drought risk quantification.

Synthetic-aperture radar (SAR) instruments offer unique features to image vegetation and crop cover due to their all-weather capabilities and ability to capture vegetation characteristics different from those derived from optical instruments. Due to this, SAR imagery can be captured at the best suited dates, and the probability of having close-to-perfect SAR time series is much higher than having optimal optical time series.

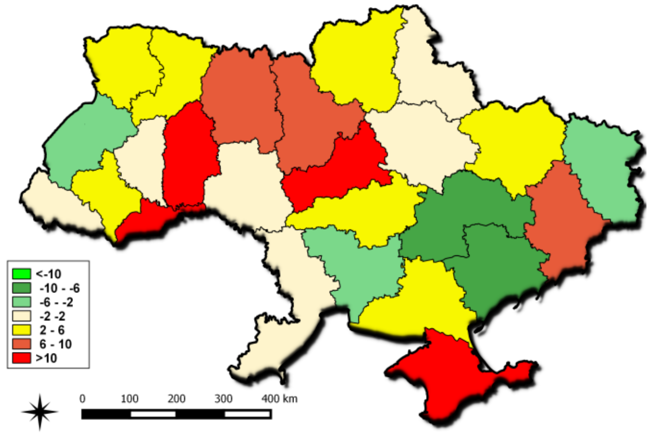

Fig. 2. Land use changes since 1990 till 2010 averaged by administrative regions.

Using this data and historical data from the past 3 years we estimated the areas of the most important industrial crops and developed maps of crop rotation violations (Fig. 3). According to our results 29% of winter wheat and 48.8% of maize were grown with crop rotation violations (2 years in a row).

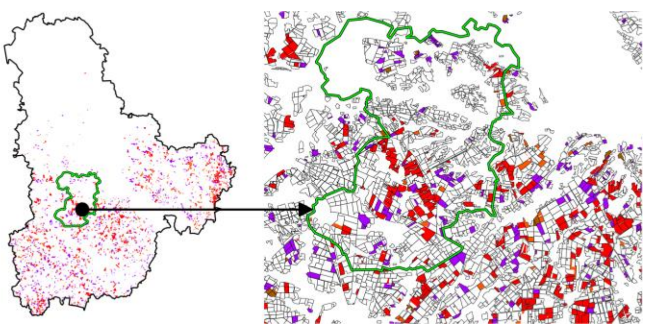

Fig. 3. Crop rotation violations for Kyiv region in 2015. Red color shows violation for maize, and purple shows crop rotation violations for winter wheat

These results are obtained within multiple international initiatives including Global Agriculture Monitoring System (GEO-GLAM) [7], Joint Experiment of Crop Assessment and Monitoring (JECAM) [8], and FP7 projects "Stimulating Innovation for Global Monitoring of Agriculture and its Impact on the Environment in support of GEOGLAM" (SIGMA, www.geoglam-sigma.info) and "Implementation of Multi-scale Agricultural Indicators Exploiting Sentinels" (ImagineS, http://fp7-imagines.eu/). Developed method is tested for multiple test sites across wide range of agricultural systems. Proposed approach can be used for risk assessment [15]-[17], when the risk can be regarded as a probabilistic prediction of a current situation or future hazardous events and associated losses.

4. DISCUSSIONS AND STEPS FORWARD

Our results are fully aligned with the strategic objectives of existing international initiatives and programs. With the developed approach we are going to contribute to the following international programs.

GEO/GEOSS. We are planning to participate in "Strategic Plan 2016-2025: Implementing GEOSS", particularly to the following areas: "Disaster Resilience" and "Food Security and Sustainable Agriculture".

Copernicus. Due to developed automatic chain of "big data" processing we are able to heavily utilize satellite data provided within the European Copernicus program [18] - [20]. In particular, the added-value and potential innovative services can be developed from Sentinel-1 and Sentinel-2 imagery in synergy with other satellites (Landsat-8) and further valorize exploitation of these datasets.

Climate Change Initiative. Our further plans are related to the Earth Observation programs (i.e. the Living Planet Programme, the European Earth Watch Programme - Global Monitoring of Essential Climate Variables known as the "Climate Change Initiative") of the European Space Agency (ESA) by using satellite data for improving the prediction of the spatial and temporal distribution of disasters and hence mitigating their consequences for society.

Horizon2020 ERA-NET project. The results are in the line with the newly accepted project "The European Network for Observing Our Changing Planet" which aims at developing a Transnational Environmental Observation System in Support of European & International Policies.

So, the developed methodology opens the prospects to contribute to many important international programs.

5. REFERENCES

[1] Y. Ma, H. Wu, L. Wang, B. Huang, R. Ranjan, A. Zomaya, and W. Jie, "Remote sensing big data computing: Challenges and opportunities," *Future Generation Computer Systems*, 51, pp. 47–60, 2015.

[2] Y. Ma, L. Wang, P. Liu, and R. Ranjan, "Towards building a data-intensive index for big data computing – A case study of Remote Sensing data processing," *Information Sciences*, 319, pp. 171–188, 2015.

[3] N. Kussul, A. Shelestov, R. Basarab, S. Skakun, O. Kussul, and M. Lavreniuk, "Geospatial intelligence and data fusion techniques for sustainable development problems," *CEUR Workshop Proceedings*, 1356, pp. 196–203, 2015.

[4] C. A. Mattmann, "Computing: A vision for data science," *Nature*, 493 (7433), pp. 473–475, 2013.

[5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, 521 (7553), pp. 436–444, 2015.

[6] M. Lavreniuk, N. Kussul, S. Skakun, A. Shelestov, and B. Yailymov, “Regional retrospective high resolution land cover for Ukraine: methodology and results,” In *2015 IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), pp. 3965–3968, 2015.

[7] I. Becker-Reshef, C. Justice, M. Sullivan, E. Vermote, C. Tucker, A. Anyamba, J. Small, E. Pak, E. Masuoka, J. Schmaltz, M. Hansen, K. Pittman, C. Birkett, D. Williams, C. Reynolds, and B. Doorn, “Monitoring Global Croplands with Coarse Resolution Earth Observations: The Global Agriculture Monitoring (GLAM) Project,” *Remote Sensing*, vol. 2, pp. 1589–1609, 2010.

[8] N. Kussul, S. Skakun, A. Shelestov, and O. Kussul, “The use of satellite SAR imagery to crop classification in Ukraine within JECAM project,” in: IGARSS 2014, 13-18 July 2014, pp. 1497–1500, 2014.

[9] N. Kussul, S. Skakun, A. Shelestov, O. Kravchenko, J.F. Gallego, and O. Kussul, “Crop area estimation in Ukraine using satellite data within the MARS project,” in: IGARSS 2012, 22-27 July 2012, Munich, Germany, pp. 3756–3759.

[10] R. Agarwal, P. Ranjan, and H. Chipman, “A new Bayesian ensemble of trees approach for land cover classification of satellite imagery,” *Canadian Journal of Remote Sensing*, vol. 39, no. 6, pp. 507–520, 2014.

[11] S. Skakun, N. Kussul, A. Shelestov, M. Lavreniuk, and O. Kussul, “Efficiency Assessment of Multitemporal C-Band Radarsat-2 Intensity and Landsat-8 Surface Reflectance Satellite Imagery for Crop Classification in Ukraine,” in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, doi: 10.1109/JSTARS.2015.2454297.

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

[13] N. Kussul, G. Lemoine, J. Gallego, S. Skakun, and M. Lavreniuk, “Parcel based classification for agricultural mapping and monitoring using multi-temporal satellite image sequences,” In *2015 IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), pp. 165–168, 2015.

[14] F.J. Gallego, N. Kussul, S. Skakun, O. Kravchenko, A. Shelestov, and O. Kussul, “Efficiency assessment of using satellite data for crop area estimation in Ukraine,” *Int. J. Appl. Earth Observ. Geoinf.*, vol. 29, pp. 22-30, 2014.

[15] D. Guha-Sapir, F. Vos, R. Below, and S. Ponserre, “Annual Disaster Statistical Review 2011: The Numbers and Trends,” Brussels: Centre for Research on the Epidemiology of Disasters (CRED), 2012.

[16] S. Skakun, N. Kussul, A. Shelestov, and O. Kussul, “The use of satellite data for agriculture drought risk quantification in Ukraine,” *Geomatics, Natural Hazards and Risk*, 2015, doi: 10.1080/19475705.2015.1016555.

[17] S. Skakun, N. Kussul, A. Shelestov, and O. Kussul, “Flood Hazard and Flood Risk Assessment Using a Time Series of Satellite Images: A Case Study in Namibia,” *Risk Anal.*, vol. 34, no. 8, pp. 1521–1537, 2014.

[18] A.Yu. Shelestov, A.N. Kravchenko, S.V Skakun, S.V. Voloshin, and N.N. Kussul, “Geospatial information system for agricultural monitoring,” *Cybernetics Syst. Anal.*, vol. 49, no. 1, pp. 124–132, 2013.

[19] N. Kussul, A. Shelestov and S. Skakun, “Grid and sensor web technologies for environmental monitoring,” *Earth Science Informatics*, vol. 2, no. 1-2, pp. 37–51, 2009.

[20] A. Shelestov, N. Kussul, and S. Skakun, “Grid technologies in monitoring systems based on satellite data,” *Journal of Automation and Information Sciences*, vol. 38, no. 3, pp. 69–80, 2006.