

Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data

Natalia Kussul, Mykola Lavreniuk, Sergii Skakun, and Andrii Shelestov

Abstract—Deep learning (DL) is a powerful state-of-the-art technique for image processing including remote sensing (RS) images. This letter describes a multilevel DL architecture that targets land cover and crop type classification from multitemporal multisource satellite imagery. The pillars of the architecture are unsupervised neural network (NN) that is used for optical imagery segmentation and missing data restoration due to clouds and shadows, and an ensemble of supervised NNs. As basic supervised NN architecture, we use a traditional fully connected multilayer perceptron (MLP) and the most commonly used approach in RS community random forest, and compare them with convolutional NNs (CNNs). Experiments are carried out for the joint experiment of crop assessment and monitoring test site in Ukraine for classification of crops in a heterogeneous environment using nineteen multitemporal scenes acquired by Landsat-8 and Sentinel-1A RS satellites. The architecture with an ensemble of CNNs outperforms the one with MLPs allowing us to better discriminate certain summer crop types, in particular maize and soybeans, and yielding the target accuracies more than 85% for all major crops (wheat, maize, sunflower, soybeans, and sugar beet).

Index Terms—Agriculture, convolutional neural networks (CNNs), crop classification, deep learning (DL), joint experiment of crop assessment and monitoring (JECAM), Landsat-8, remote sensing (RS), Sentinel-1, TensorFlow, Ukraine.

I. INTRODUCTION

THE last several years and onward could be called the years of Big Free Data in remote sensing (RS). During the 2013–2016 period, several optical and synthetic aperture radar (SAR) RS satellites were launched with high spatial resolution (10–30 m), in particular Sentinel-1A/B and Sentinel-2A within the European Copernicus program [1], [2], and Landsat-8 within the Landsat Project, a joint initiative between the U.S. Geological Survey (USGS) and the National Aeronautics and Space Administration [3]. These data sets are freely available on operational basis. This opens unprecedented opportunities for a wide range of preoperational and operational applications in the environment and agricultural domains taking advantage of high temporal resolution data sets and advances in the

Manuscript received February 17, 2017; accepted March 6, 2017. Date of publication March 31, 2017; date of current version April 20, 2017.

N. Kussul and M. Lavreniuk are with the Department of Space Information Technologies and Systems, Space Research Institute, National Academy of Sciences of Ukraine and SSA Ukraine, 03680 Kyiv, Ukraine (e-mail: inform@ikk.kiev.ua; natalia.kussul@gmail.com; nick_93@ukr.net).

S. Skakun is with the Department of Geographical Sciences, University of Maryland, College Park, MD 20742 USA (e-mail: skakun@umd.edu).

A. Shelestov is with the Department of Information Security, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute,” 03056 Kyiv, Ukraine (e-mail: andrii.shelestov@gmail.com).

Color versions of one or more of the figures in this letter are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LGRS.2017.2681128

multisources data fusion techniques [4], [5]. Land cover and crop type maps are one of the most essential inputs when dealing with environmental and agriculture monitoring tasks [6]–[8]. Multitemporal multisource satellite imagery is usually required in order to capture specific crop growth stages and thus being able to discriminate different crop types. For example, multispectral optical imagery only might not be enough to discriminate summer crops in a complex and heterogeneous environment. For this, SAR-derived information adds an added value that allows discrimination of particular crop types [9], [10].

A comprehensive study on the state-of-the-art supervised pixel-based methods for land cover mapping was performed by Khatami *et al.* [11]. They found that support vector machine (SVM) was the most efficient for most applications with an overall accuracy (OA) of about 75%. The second method with approximately the same efficiency (74% of OA) was a neural network (NN)-based classifier. In that study, classification was done only for a single date image. At the same time, SVM is too much resource consuming to be used for big data applications and large area classification problems. Another popular approach in the RS domain is the random forest (RF)-based approach [12]. However, multiple features should be engineered to feed the RF classifier for the efficient use.

Over the past few years, the most popular and efficient approaches for multisensor and multitemporal land cover classification are ensemble-based [13]–[16] and deep learning (DL) [17]–[20]. These techniques are found to outperform the SVM [21]–[23]. DL is a powerful machine learning methodology for solving a wide range of tasks arising in image processing, computer vision, signal processing, and natural language processing [24]. The main idea is to simulate the human vision to deal with big data problem, use all the data available and provide the semantic information at the output. Plenty of models, frameworks and benchmark databases of reference imagery are available for image classification domain. Over past years, more and more studies have been using DL for processing of RS imagery [25], [26]. DL proved to be efficient for processing both optical (hyperspectral and multispectral imagery) and radar images, in extracting different land cover types such as road extraction, buildings extraction [17], [27], [28]. In terms of particular DL architectures, convolutional NNs (CNNs), deep autoencoders, deep belief networks, and recurrent NN with long short-term memory model have already been explored for RS tasks [17], [28]–[31]. It should be noted that most studies with DL for RS utilize a single date image for classification purposes, e.g., land

TABLE I
DATES OF ACQUISITION OF LANDSAT-8 AND SENTINEL-1A
FOR THE KYIV REGION IN 2015

Region and Year	Landsat-8	Sentinel-1
Kyiv 2015	24.05, 9.06, 25.06, 28.08 (181/24, 181/25, 181/26)	01.03, 13.03, 25.03, 06.04, 18.04, 30.04,
		12.05, 24.05, 05.06,
		17.06, 29.06, 11.07,
		23.07, 16.08, 28.08

cover or object detection. However, multitemporal images are usually required to reliably identify specific land cover classes such as crop types.

When providing large scale crop mapping using multitemporal 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 spaceborne SAR imagery is characterized by backscatter intensity and phase in multiple polarizations. Both of the data sources have multitemporal nature and different spatial resolutions. That is why, DL implementation for land cover and crops classification based on data fusion of multitemporal multisensor satellite data is a challenge.

In this letter, we propose a multilayer DL architecture that is targeted for classification of multisource multitemporal RS images, both optical and SAR, at a pixel level. The core of the architecture is an ensemble of CNNs. The proposed architecture is applied for crop classification using Landsat-8 and Sentinel-1A time-series and provides accuracy high enough (>85%) to be considered for operational context at the national level [9].

II. STUDY AREA AND MATERIALS

We address the problem of land cover and crop classification for Kyiv region of Ukraine using multitemporal multisource images acquired by Landsat-8 and Sentinel-1A satellites. 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. Such technology is particularly important taking into account national level demonstration within the ESA “Sentinel 2 for Agriculture” project (Sen2Agri), started in 2015.

For the 2015 vegetation season (since October 2014 till September 2015) four Landsat-8 and fifteen Sentinel-1 images were acquired for the study area (Table I). Atmospherically corrected Landsat-8 images downloaded from the USGS earth-Explorer system were used in this letter [32]. The Landsat-8 product is provided with cloud and shadow masks [33]. These regions were masked as areas without data. Sentinel-1A images went through a preprocessing procedure that cov-

TABLE II
NUMBER AND AREA OF POLYGONS COLLECTED DURING
GROUND SURVEYS FOR THE KYIV REGION IN 2015

#	Growing period	Class	Fields		Area, ha
			#	%	
1	End of September – mid of July	Winter wheat	102	18.6	3695.9 19.5
2	End of August – mid of July	Winter rapeseed	22	4	715.9 3.8
3	End of April – mid of July	Spring crops	11	2	296 1.6
4	End of April – end of August	Maize	98	17.9	4329.1 22.8
5	Start of May – end of October	Sugar beet	8	1.5	860.7 4.5
6	Start of May – start of September	Sunflower	53	9.7	1954 10.3
7	Start of May – end of September	Soybeans	87	15.9	3006.9 15.8
8	—	Forest	49	9	2012.3 10.6
9	—	Grassland	64	11.7	952.3 5
10	—	Bare land	10	1.8	71.4 0.4
11	—	Water	43	7.9	1072.1 5.7
Total			547		18966.6

ered calibration, multilooking (with 2×2 window), speckle filtering (3×3 window with Refined Lee algorithm), and terrain correction using The Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM). A time-series of six spectral bands from each Landsat-8 scene and two bands with combinations of vertical transmit and vertical receive and combinations of vertical transmit and horizontal receive polarizations from each Sentinel-1 scene is used as an input to the classification model. In the study for the 2015 vegetation season, we used four images acquired by Landsat-8 and 15 images acquired by Sentinel-1A. Ground truth data were collected during ground surveys in May–July of 2015 to generate training and testing sets to train and validate the proposed classifier, respectively. Ground surveys were conducted along the road as an adopted sampling design for the joint experiment of crop assessment and monitoring (JECAM) experiments [34]. In total, 547 polygons of different classes were collected (Table II). These polygons were randomly divided into training (calibration) set (50%) and validation set (50%).

III. METHODOLOGY

A. General Architecture Overview

A four-level architecture is proposed for classification of crop types from multitemporal satellite imagery. These levels are preprocessing, supervised classification, postprocessing, and geospatial analysis (Fig. 1).

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 preprocessing step should be performed to 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 letter. We propose different CNNs architectures, namely, 1-D and 2-D, to explore spectral and spatial features, respectively. To our best knowledge, this is the first attempt to apply CNNs to multisource multitemporal satellite imagery for crop classification. The CNNs architecture is compared to the existing methods such as an ensemble of multilayer perceptrons (MLPs) (ENN) and RF classifier. *Levels III* and *IV*

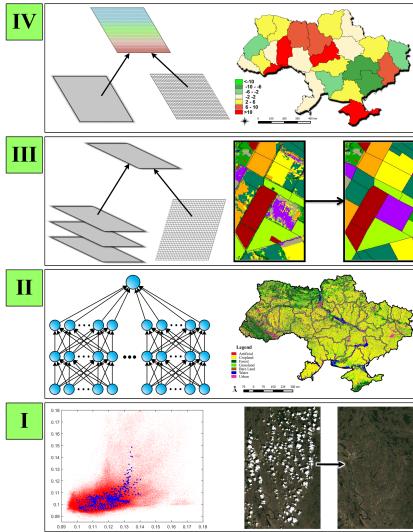


Fig. 1. Four-level hierarchical DL model for satellite data classification and land cover/land use changes analysis (I—preprocessing for dealing with missing data on optical images due to clouds/shadows; II—supervised classification; III—postprocessing using additional geospatial data to improve classification maps; IV—geospatial analysis for a high-level product, e.g., crop area estimation).

are aimed at improving the resulting classification map with available geospatial layers and building high-level products. The latter can be crop area estimation and crop rotation area estimation. All these levels of the architecture are described in more detail in the following sections.

B. Level I: Preprocessing

For preprocessing, we utilize self-organizing Kohonen maps (SOMs) for optical images segmentation and subsequent restoration of missing data in a time-series of satellite imagery [35]. SOMs are trained for each spectral band separately using nonmissing 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 postprocessing procedure (at *level III*) to improve the resulting classification map [16]. The detailed description of the restoration algorithm is given in [35] and [36].

C. Level II: Supervised Classification

1) General Overview: The core element of the model is the supervised classification, which is performed at the second stage (*level II*). We explore two different paradigms: state-of-the-art methods (RF an ENN) and compare those classifiers with proposed ensemble of CNNs. Each MLP represents a classical fully connected NN with a single hidden layer. The MLP transforms an input into a feature space by a hidden layer, and features are subsequently used to discriminate classes by the output layer. The CNN, in turn, builds a hierarchical set of features through local convolution and down-sampling. The resulting feature maps are used to discriminate

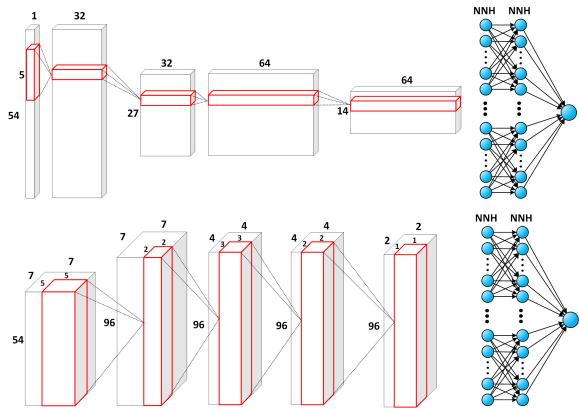


Fig. 2. Deep CNN architecture. (Top) 1-D CNN architecture. (Bottom) 2-D CNN architecture.

classes with fully connected output layer. The detailed description of the ENN approach is given in [34] and [36], while the CNN architecture is described in the next section. We used the Orfeo Toolbox implementation for the RF classifier.

2) Supervised Classification With Convolutional Neural Networks: The two bands from each of the fifteen Sentinel-1A scenes and the six bands from each of the four Landsat-8 scenes form a CNN input feature vector with dimension size 54 ($15 \times 2 + 4 \times 6$). 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 pixel but also small objects like roads, and forest “stripes” and clear cuts within the forest (with linear dimensions of several pixels) are missed. In this letter, we compare two different CNN architectures: 1-D CNN with convolutions in the spectral domain [37], and 2-D CNN with convolutions in the spatial domain. 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. 2). We used a rectified linear unit (ReLU) function that is one of the most popular and efficient activation functions for deep NNs. 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. In both architectures, there are five CNNs in the ensemble. 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 being 60, 70, 80, 90, and 100 for five CNNs, respectively.

Tikhonov's L2 regularization, dropout with probability of 0.5 and learning rate exponential decay techniques are used to prevent the overfitting problem and to generalize the loss function. For loss function optimization, we used advanced adaptive moment estimation method that is a combination of AdaGrad and RMSProp methods and has faster convergence comparing to the well-known methods such as gradient descent, stochastic gradient descent, AdaGrad and RMSProp [38]. Batch learning technique with sample size 32 is used to speed-up the NN training phase. Multiclass cross

entropy function is used as a loss function and softmax function is used to provide the *a posteriori* probability for each class. Ensembles of 1-D and 2-D CNNs are implemented using the Google's library TensorFlow [39].

The proposed 1-D CNN architecture is able to provide classification for each pixel of the input image. In turn, a classical 2-D CNN provides a class for a window with size of 7×7 pixels. In this case, there is a popular approach with a direct up-scaling to match the input size that will actually reduce the spatial resolution of the classification map (at the order of four times for two down-scaling layers with filter size 2×2). In our experiment, we utilized a sliding window technology with a 1 pixel step and the resulting class was assigned to the central pixel of the sliding window. Therefore, the output classification map had the same spatial resolution, but even in this case some small objects were smoothed and misclassified.

D. Level III and IV: Postprocessing and Geospatial Analysis

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, for example parcels [16]. 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 multisourced 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 [40].

IV. RESULTS

Overall classification accuracies for RF, ENN, ensemble of 1-D and 2-D CNNs were 88.7%, 92.7%, 93.5%, and 94.6%, respectively (Table III). User's and producer's accuracies (UA and PA) provided by an ensemble of 2-D CNNs were the highest for all classes. The RF classifier provided the lowest accuracies comparing to NN-based approaches. Accuracy for winter rapeseed, spring crops, sunflower, forest, and water did not vary significantly with different approaches. At the same time, major improvements of using CNNs comparing to RF were achieved for maize, sugar beet, soybeans, grassland, and bare land (Fig. 3). Usually, the main confusion in crop classification map for Ukraine territory is confusion between maize and soybeans. Using the ensemble of 2-D CNNs, we were able to discriminate these classes more reliably: maize (PA = 94.6%, UA = 93.6%) and soybeans (PA = 86.9%, UA = 89.1%).

All these experiments were executed on a computer with Intel Core i7-4770 processor and RAM 32 Gb. Training of ensemble of MLPs took up to 10 min at the same time ensemble of 1-D CNNs trained approximately 4 h and 2-D CNNs training takes about 12 h.

V. CONCLUSION

In this letter, we proposed a multilevel DL approach for land cover and crop types classification using multitemporal

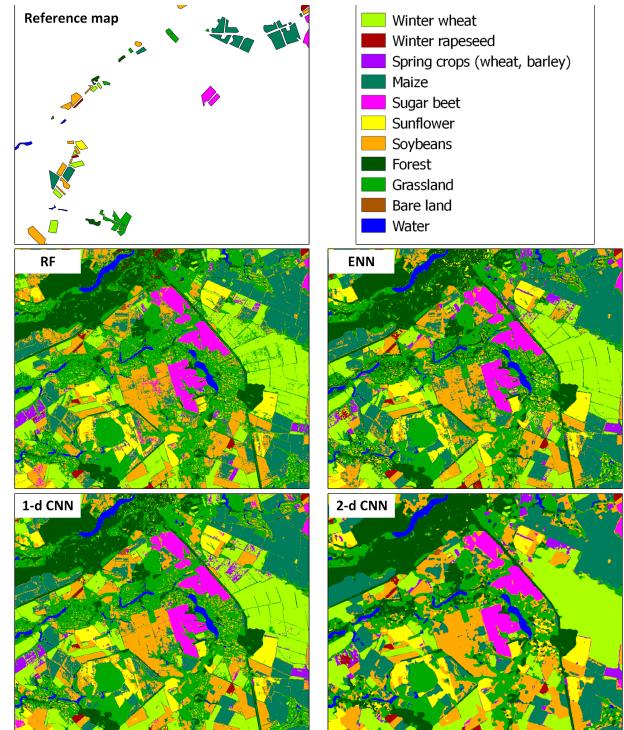


Fig. 3. Example of classification result for the Kyiv region for 2015 based on all Landsat-8 and Sentinel-1A images.

TABLE III
COMPARISON OF PA, UA, AND OA FOR THE RF, ENN,
1-D CNNs, AND 2-D CNNs FOR KYIV REGION IN 2015

#	Class	RF		ENN		1-d CNNs		2-d CNNs	
		PA	UA	PA	UA	PA	UA	PA	UA
1	Winter wheat	92.2	95.8	95.2	95.3	95.1	95.9	95.3	96.1
2	Winter rapeseed	97.3	78.2	97.5	78.4	97.1	78.3	98.7	78.6
3	Spring crops	49.1	96.1	50.6	96.7	48.6	93.6	51.1	97
4	Maize	79.3	88.2	90.6	90.8	90.6	94.1	94.6	93.6
5	Sugar beet	91.3	95.7	97.5	98.4	97.4	99.1	98.4	100
6	Sunflower	98.9	96.5	99.2	96.3	99	96.4	99.4	96.9
7	Soybeans	79	69.8	81	83.9	87.3	84.5	86.9	89.1
8	Forest	99.5	99.5	99.6	99.7	99.4	99.8	99.4	99.9
9	Grassland	92	82	94.5	91.6	94.6	91.8	93.5	94.5
10	Bare land	89.6	47.4	90.6	82.8	92.6	74.9	89.3	88.2
11	Water	99.8	100	99.8	99.9	99.9	100	99.9	100
OA, %		88.7		92.7		93.5		94.6	

multisource satellite imagery. The architecture uses both unsupervised and supervised NNs for segmentation and subsequent classification of satellite imagery, respectively. In this letter, we used Landsat-8 and Sentinel-1A images over the JECAM test site in Ukraine. Ensemble of 1-D and 2-D CNNs outperformed the RF classifier and an ensemble of MLPs allowing us to better discriminate summer crops, in particular maize and soybeans. In general, the use of CNN allowed us to reach the target accuracy of 85% for major crops (wheat, maize, sunflower, soybeans, and sugar beet) thus making a foundation for further operational use of RS data for the whole territory of Ukraine within the Sentinel-2 for Agriculture project. The main advantage of using CNNs over MLP and RF is that it enables to build a hierarchy of local and sparse features derived from spectral and temporal profiles while MLP and

RF build a global transformation of features. The 2-D CNNs outperformed the 1-D CNNs, but some small objects in the final classification map provided by 2-D CNNs were smoothed and misclassified.

REFERENCES

- [1] M. Drusch *et al.*, "Sentinel-2: ESA's optical high-resolution mission for GMES operational services," *Remote Sens. Environ.*, vol. 120, pp. 25–36, May 2012.
- [2] R. Torres *et al.*, "GMES Sentinel-1 mission," *Remote Sens. Environ.*, vol. 120, pp. 9–24, May 2012.
- [3] D. P. Roy *et al.*, "Landsat-8: Science and product vision for terrestrial global change research," *Remote Sens. Environ.*, vol. 145, pp. 154–172, Apr. 2014.
- [4] J. Zhang, "Multi-source remote sensing data fusion: Status and trends," *Int. J. Image Data Fusion*, vol. 1, no. 1, pp. 5–24, Nov. 2010.
- [5] M. D. Mura, S. Prasad, F. Pacifici, P. Gamba, J. Chanussot, and J. A. Benediktsson, "Challenges and opportunities of multimodality and data fusion in remote sensing," *Proc. IEEE*, vol. 103, no. 9, pp. 1585–1601, Sep. 2015.
- [6] A. Kolotii *et al.*, "Comparison of biophysical and satellite predictors for wheat yield forecasting in Ukraine," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 40, no. 7, p. 35, 2015, doi: 10.5194/isprsarchives-XL-7-W3-39-2015.
- [7] F. Kogan *et al.*, "Winter wheat yield forecasting: A comparative analysis of results of regression and biophysical models," *J. Autom. Inf. Sci.*, vol. 45, no. 6, pp. 68–81, 2013.
- [8] F. Kogan *et al.*, "Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 23, pp. 192–203, Aug. 2013.
- [9] H. McNairn, A. Kross, D. Lapen, R. Caves, and J. Shang, "Early season monitoring of corn and soybeans with TerraSAR-X and RADARSAT-2," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 28, pp. 252–259, May 2014.
- [10] S. Skakun, N. Kussul, A. Y. 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," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 8, pp. 3712–3719, Aug. 2016.
- [11] R. Khatami, G. Mountakis, and S. V. Stehman, "A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research," *Remote Sens. Environ.*, vol. 177, pp. 89–100, May 2016.
- [12] P. O. Gislason, J. A. Benediktsson, and J. R. Sveinsson, "Random Forests for land cover classification," *Pattern Recognit. Lett.*, vol. 27, no. 4, pp. 294–300, 2006.
- [13] M. Han, X. Zhu, and W. Yao, "Remote sensing image classification based on neural network ensemble algorithm," *Neurocomputing*, vol. 78, no. 1, pp. 133–138, 2012.
- [14] X. Huang and L. Zhang, "An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 257–272, Jan. 2013.
- [15] M. S. Lavreniuk *et al.*, "Large-scale classification of land cover using retrospective satellite data," *Cybern. Syst. Anal.*, vol. 52, no. 1, pp. 127–138, 2016.
- [16] N. Kussul, G. Lemoine, F. J. Gallego, S. V. Skakun, M. Lavreniuk, and A. Y. Shelestov, "Parcel-based crop classification in Ukraine using Landsat-8 data and Sentinel-1A data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 6, pp. 2500–2508, Jan. 2016.
- [17] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.
- [18] W. Zhao and S. Du, "Learning multiscale and deep representations for classifying remotely sensed imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 113, pp. 155–165, Mar. 2016.
- [19] N. Kussul, N. Lavreniuk, A. Shelestov, B. Yailymov, and I. Butko, "Land cover changes analysis based on deep machine learning technique," *J. Autom. Inf. Sci.*, vol. 48, no. 5, pp. 42–54, 2016.
- [20] N. Kussul, A. Shelestov, R. Basarab, S. Skakun, O. Kussul, and M. Lavreniuk, "Geospatial intelligence and data fusion techniques for sustainable development problems," in *Proc. ICTERI*, 2015, pp. 196–203.
- [21] J. Ding, B. Chen, H. Liu, and M. Huang, "Convolutional neural network with data augmentation for SAR target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 364–368, Mar. 2016.
- [22] F. J. Huang and Y. LeCun, "Large-scale learning with SVM and convolutional for generic object categorization," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2006, pp. 284–291.
- [23] T. Ishii, R. Nakamura, H. Nakada, Y. Mochizuki, and H. Ishikawa, "Surface object recognition with CNN and SVM in Landsat 8 images," in *Proc. 14th IAPR Int. Conf. Mach. Vis. Appl. (MVA)*, May 2015, pp. 341–344.
- [24] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
- [25] F. Zhang, B. Du, and L. Zhang, "Saliency-guided unsupervised feature learning for scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2175–2184, Apr. 2015.
- [26] F. Zhang, B. Du, and L. Zhang, "Scene classification via a gradient boosting random convolutional network framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1793–1802, Mar. 2016.
- [27] V. Mnih and G. E. Hinton, "Learning to detect roads in high-resolution aerial images," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 210–223.
- [28] J. Geng, J. Fan, H. Wang, X. Ma, B. Li, and F. Chen, "High-resolution SAR image classification via deep convolutional autoencoders," *IEEE Trans. Geosci. Remote Sens.*, vol. 12, no. 11, pp. 2351–2355, Nov. 2015.
- [29] Y. Chen, X. Zhao, and X. Jia, "Spectral-spatial classification of hyperspectral data based on deep belief network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2381–2392, Jun. 2015.
- [30] H. Liang and Q. Li, "Hyperspectral imagery classification using sparse representations of convolutional neural network features," *Remote Sens.*, vol. 8, no. 2, p. 99, 2016.
- [31] H. Lyu, H. Lu, and L. Mou, "Learning a transferable change rule from a recurrent neural network for land cover change detection," *Remote Sens.*, vol. 8, no. 6, p. 506, 2016.
- [32] E. Vermote, C. Justice, M. Claverie, and B. Franch, "Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product," *Remote Sens. Environ.*, vol. 185, pp. 46–56, 2016, doi: 10.1016/j.rse.2016.04.008.
- [33] Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images," *Remote Sens. Environ.*, vol. 159, pp. 269–277, Mar. 2015.
- [34] F. Waldner *et al.*, "Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity," *Int. J. Remote Sens.*, vol. 37, no. 14, pp. 3196–3231, 2016.
- [35] S. V. Skakun and R. M. Basarab, "Reconstruction of missing data in time-series of optical satellite images using self-organizing Kohonen maps," *J. Autom. Inform. Sci.*, vol. 46, no. 12, pp. 19–26, 2014.
- [36] N. Kussul, S. Skakun, A. Shelestov, M. Lavreniuk, B. Yailymov, and O. Kussul, "Regional scale crop mapping using multi-temporal satellite imagery," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 40, no. 7, pp. 45–52, 2015.
- [37] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification," *J. Sens.*, vol. 2015, art. no. 258619, 2015.
- [38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization." Unpublished paper, 2014. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [39] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous distributed systems." Unpublished paper, 2016. [Online]. Available: <https://arxiv.org/abs/1603.04467>
- [40] 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 Observat. Geoinf.*, vol. 29, pp. 22–30, Jun. 2014.