IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium | 978-1-6654-2792-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/IGARSS46834.2022.9883655

AN AUTOMATIC APPROACH FOR THE PRODUCTION OF A TIME SERIES OF CONSISTENT LAND-COVER MAPS BASED ON LONG-SHORT TERM MEMORY

Rocco Sedona[1,2], *Claudia Paris*[3], *Liang Tian*[2], *Morris Riedel*[1,2], *Gabriele Cavallaro*[1]

[1] Jülich Supercomputing Centre, Forschungszentrum Jülich, Germany
[2] School of Engineering and Natural Sciences, University of Iceland, Iceland
[3] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, The Netherlands

ABSTRACT

This paper presents an approach that aims to produce a Time-Series (TS) of consistent Land-Cover (LC) maps, typically needed to perform environmental monitoring. First, it creates an annual training set for each TS to be classified, leveraging on publicly available thematic products. These annual training sets are then used to generate a set of preliminary LC maps that allow for the identification of the unchanged areas, i.e., the stable temporal component. Such areas can be used to define an informative and reliable multi-year training set, by selecting samples belonging to the different years for all the classes. The multi-year training set is finally employed to train a unique multi-year Long Short Term Memory (LSTM) model, which enhances the consistency of the annual LC maps. The preliminary results carried out on three TSs of Sentinel 2 images acquired in Italy in 2018, 2019 and 2020 demonstrates the capability of the method to improve the consistency of the annual LC maps. The agreement of the obtained maps is $\approx 78\%$, compared to the $\approx 74\%$ achieved by the LSTM models trained separately.

***Index Terms*—** Deep Learning (DL) Models, Long Short Term Memory (LSTM), Time-Series (TS) of Consistent Land-Cover (LC) Maps, Multi-year training set.

1. INTRODUCTION

The dense Time-Series (TS) of images with a worldwide coverage provided by Sentinel 1 and Sentinel 2 allow the production of large-scale Land-Cover (LC) maps in a timely manner [1]. For this reasons, several methods have been recently proposed to produce maps at country, continental or global scale [2, 3]. However, when the maps are produced separately there is the risk of showing unrealistic year-to-year LC changes. The map consistency is crucial when monitoring complex environmental processes such as desertification, arctic greening or soil erosion. While a lot of effort has been devoted to generating annual maps, little has been done for the production of consistent thematic products.

In [4] a method for mapping global LC types from 2001 to 2010 at 250 m resolution with multiple year TSs of MODIS data is proposed. The strategy is to generate a map for each single year by using the data acquired in the preceding and subsequent years as well. In [5], the authors propose to apply a Hidden Markov Model (HMM) as a post-processing step to a TS of LC maps to help distinguish real LC change from spurious changes arising from errors in classification. On the one hand, these methods have demonstrated improvements on the temporal consistency of classification maps. On the other hand, these strategies may lead to the risk of losing inter-annual LC changes, especially when the analysis includes long TSs of data.

In this paper, we propose a novel approach which aims to produce a TS of consistent annual LC products based on the Long Short Term Memory (LSTM) multitemporal Deep Learning (DL) model. Contrary to the above-mentioned approaches, our method uses only the TS of the year under study and does not impose constraints with the application of the post-processing analysis. Furthermore, instead of separately classifying the TSs of Earth Observation (EO) data acquired in different years, our method trains one LSTM model with the multi-year training set to produce a TS of consistent LC maps. First, it extracts a training set per year leveraging on publicly available thematic product. The annual training sets are used to separately train different Random Forest (RF) models to detect unchanged area, which can be used to produce a reliable multi-year training set. Finally, the multi-year LSTM model is trained to produce a set of annual LC maps. The advantages of the proposed approach are: (1) the automatic production of a multi-year training set, (2) the use of a LSTM for capturing the temporal trends, and (3) the training of a unique LSTM model using multi-year TSs of EO data, which enhances the consistency of the annual LC maps.

Part of this work was performed in the CoE RAISE project which has received funding from the European Union's Horizon 2020 Research and Innovation Framework Programme H2020-INFRAEDI-2019-1 under grant agreement no. 951733. The authors gratefully acknowledge the computing resources from the DEEP-EST project, which received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement no. 754304.

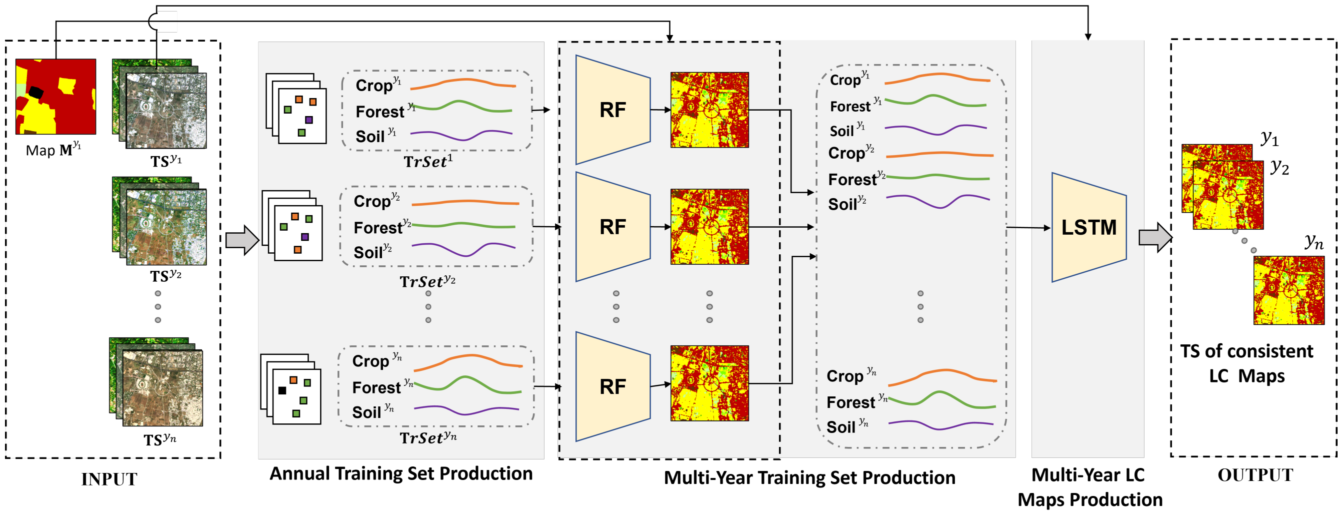

Fig. 1: Flow chart of the proposed approach which aims to produce a TS of consistent LC maps.

2. PROPOSED MULTI-YEAR MAPPING METHOD

Figure 1 shows the block-scheme of the proposed method made up of three main steps: (i) the annual training set production, (ii) the multi-year training set production, and (iii) the multi-year LC maps production.

2.1. Problem Formulation and Notation

In this section we formalize the multi-year LC mapping problem and define the notation used in the paper. Let $\mathbf{TS}^{y_i} = (\mathbf{X}_1^{y_i}, \mathbf{X}_2^{y_i}, \cdots, \mathbf{X}_q^{y_i})$ be the TS made up of the q images acquired in the ith year, where $\mathbf{X}_j^{y_i} \in \mathbb{R}^{m \times n \times b}$ is a multispectral image having $m \times n$ pixels and b spectral channels, with $j = [1, \cdots, q]$ and $i = [1, \cdots, k]$. The proposed method assumes that the k TSs are atmospherically corrected, are made up of the same number of q images, and are consistent from the temporal view point (the acquisition dates of the images in all the TSs are the same) [6]. Let $\mathbf{M}^{y_1} \in \mathbb{R}^{m \times n}$ be a publicly available thematic product contemporary to one TS of EO data considered and having LC classes $\Omega = \{\omega_u\}_{u=1}^{U}$. Here for simplicity we assume that the map is contemporary to the first TS, i.e., $\mathbf{TS}^{y_1}$, however this is not a strict requirement. The map is assumed to be co-registered to the EO data and to have the same spatial resolution.

The goal of the method is to generate a TS of consistent annual LC maps $\{\hat{\mathbf{M}}^{y_1}, \hat{\mathbf{M}}^{y_2}, \cdots, \hat{\mathbf{M}}^{y_k}\}$ leveraging on: (i) the publicly available thematic product $\mathbf{M}^{y_1}$ to support the production of a multi-year training set, (ii) the temporal correlation existing between multi-year TSs acquired in the same study area $\{\mathbf{TS}^{y_1}, \mathbf{TS}^{y_2}, \cdots, \mathbf{TS}^{y_k}\}$ and, (iii) the capability of the LSTM network to capture the temporal dynamic. Please note that the proposed method can be applied to any EO data without geographical constrains due to the availability of many global LC maps.

2.2. Annual Training Set Production

This step aims to automatically generate an annual training set for each TS of EO data that have to be classified, i.e., $\{\mathbf{TS}^{y_i}\}_{i=1}^{k}$, which will be used in the next step for the production of the multi-year training set. To this end, we considered an approach similar to the one presented in [7]. The method uses the information provided by the EO data to automatically detect and extract the most reliable map labeled units. In greater detail, for each LC class ω_u, we first select all the samples in the ith $\mathbf{TS}^{y_i}$ associated to this label. Then, an automatic clustering analysis is performed to remove spurious samples not correctly associated to that label (i.e., possible changes occurred on the ground or there are classification errors present in the map). To this end, the class samples are partitioned into t_{ω_u} clusters $\{\mathbf{C}^1_{\omega_u, y_i}, \mathbf{C}^2_{\omega_u, y_i}, \ldots, \mathbf{C}^{t_{\omega_u}}_{\omega_u, y_i}\}$ according to their spectral similarity. Based on the majority decision rule, it is reasonable to assume that the dominant cluster is made up of pixels having the highest probability to be correctly associated to ω_u. The clustering is applied in a feature space made up of a set of robust spectral indices strictly connected to the physical meaning of the LC classes, i.e., Vegetation Indices (i.e., NDVI and EVI), Water Index (NDWI), Snow Index (NDSI) and Soil Index as computed in the Sen2Cor processor[1].

Once the most reliable map units are identified per class, a stratified random sampling strategy is applied to generate training sets having LC prior probabilities proportionate to those reported in $\mathbf{M}^{y_1}$. At the end of this step, the method generates a set of annual training sets $\{\mathbf{TrSet}^{y_i}\}_{i=1}^{k}$, where $\mathbf{TrSet}^{y_i} = \{(\mathbf{x}_b^{y_i}, l_b^{y_i})\}_b$ is the training set associated to the ith year having $\mathbf{x}_b^{y_i} \in \mathbb{R}^{1 \times s}$ and $l_b^{y_i} \in \Omega$, with $s = q \times b$ is the number of spectral features per number of images in the TS.

[1] https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm

2.3. Multi-year Training Set Production

The goal of this step is the production of a multi-year training set by selecting samples belonging to different years for all the classes. Such training set aims to help the LSTM model in learning the different spectral signatures of the same LC class in the different years. First, the obtained annual training sets $\{\mathbf{TrSet}^{y_i}\}_{i=1}^{k}$ are used to separately train k RF classifiers, i.e., a RF per TS. Due to the ensemble-learning strategy that combines a large set of classification trees, such models are able to deal with the presence of noise in the training set. The RF classifiers generate a set of preliminary LC maps $\{\dot{\mathbf{M}}^{y_1}, \dot{\mathbf{M}}^{y_2}, \cdots, \dot{\mathbf{M}}^{y_k}\}$ that can be used to determine the non-changed areas, i.e., where all the classified maps agree with the original thematic product $\mathbf{M}^{y_1}$. Such map agreement allows us to select the most reliable unchanged pixels across the different years. Indeed, even though the maps may also provide useful information on possible changed areas, in the considered preliminary implementation of the method we aim to include only the most reliable samples in the multi-year training set. At the end of this step, we obtain the multi-year training set, $\mathbf{TrSet} = \{(\mathbf{x}_h^{y_i}, l_h^{y_i})\}_h$, with $i = [1, \cdots, k]$ and $h \approx k \times b$.

2.4. Multi-Year LC Maps Production

The last step of the proposed method aims to produce a set of consistent TS of LC maps. To this end, we train a unique LSTM model considering the multi-year training set generated in the previous step. That model is then used to produce the TS of LC maps by classifying the corresponding annual TSs. Please note that, differently from the literature, the method does not perform any temporal smoothing step to the TS of annual LC maps. Although effective, such post-processing step may lead to the loss of changes actually occurring on the ground. For the same reason, the classification map of each year is generated considering only the TS of that specific year, instead of considering images acquired in the preceding and subsequent years. In contrast, the developed multi-year training set helps the LSTM model to better capture the different behaviours of the pixels belonging to the same class in different years, thus implicitly reinforcing the temporal consistency. In particular, the use of the same LSTM to generate the TS of LC maps allows the reduction of pixel noise across the LC maps produced for the different years, while not hampering the detection of changes occurring on the ground. At the end of the proposed approach, we obtain the set of LC maps having Ω classes will be generated, i.e., $\{\hat{\mathbf{M}}^{y_1}, \hat{\mathbf{M}}^{y_2}, \cdots, \hat{\mathbf{M}}^{y_k}\}$.

3. DATASET DESCRIPTION

In our study, we make use of acquisitions taken by the Sentinel-2 satellites. The tile that we analyze is the T32TPS, covering the area of the Trentino region, Italy. We downloaded and pre-processed 20 paired acquisitions for each of the years 2018, 2019, 2020, with a maximum difference between the date of acquisition of $\delta = 6$ days between the different years. We excluded observations with considerable cloud coverage, namely where more than 40% of the pixels are assigned to clouds in the Scene Classification Layer (SCL) map. For each year, we extracted two sets of samples randomly selected from the unchanged areas of the output maps produced by the RF classifiers. The first set is used to train the multi-year LSTM considering approximately 24.000 samples, while the second set made up of 12.000 samples was used to evaluate the loss and accuracy scores on unseen data at training time. In particular, to extract the annual training sets we relied on the CORINE Land Cover (CLC) map [8] available on the European level, considering 10 widespread LC classes, i.e., ”Artificial”, ”Grass”, ”Crops”, ”Mineral”, ”Rocks”, ”Sand”, ”Broadleaves”, ”Conifers”, ”Water”, ”Snow” .

4. EXPERIMENTAL SETUP AND RESULTS

To assess the effectiveness of the proposed approach, we compared the maps obtained with the ones generated by the LSTM separately trained per year, i.e., the standard baseline approach. Such single-year LSTM models were trained considering only the pseudo labels representing that year considering the corresponding TSs of 20 acquisitions. In contrast, the proposed approach was trained considering training samples extracted from all the TSs of images. To have a fair comparison, we considered the same LSTM model. In particular, we base our work on a *PyTorch* implementation of a LSTM with 4 layers and hidden dimension equal to 128 [2]. We trained the LSTM models for 100 epochs with the Adam optimizer on sequences of 20 observations, each with 10 features (the pixel values of the 10 considered bands), activating the dropout option. For the prediction, we split each acquisition in 25 patches of dimension 2560×2560px, and stack the features of each pixel for every observation of the year to compose the TS. We applied a median filter with kernel dimension equal to 11 to reduce the noise of the output thematic maps.

Table 1 shows the LC map agreement achieved per class and overall considering the standard baseline method, i.e., single-year LSTM and the proposed multi-year LSTM. One can notice that the proposed method is able to increase the consistency of the results obtained regardless of the LC class by increasing the overall map agreement of almost $\approx 4\%$ for both years. Figure 2 shows the number of LC changes per pixel for a portion of the considered study area when using: (a) the multi-year proposed method, (b) the single-year baseline method. Moreover, one image of 2018 and the corresponding LC map obtained with the proposed method is reported. From this result, one can notice that the proposed approach is able to reduce the classification errors at pixel level

[2]`https://github.com/dl4sits/BreizhCrops`

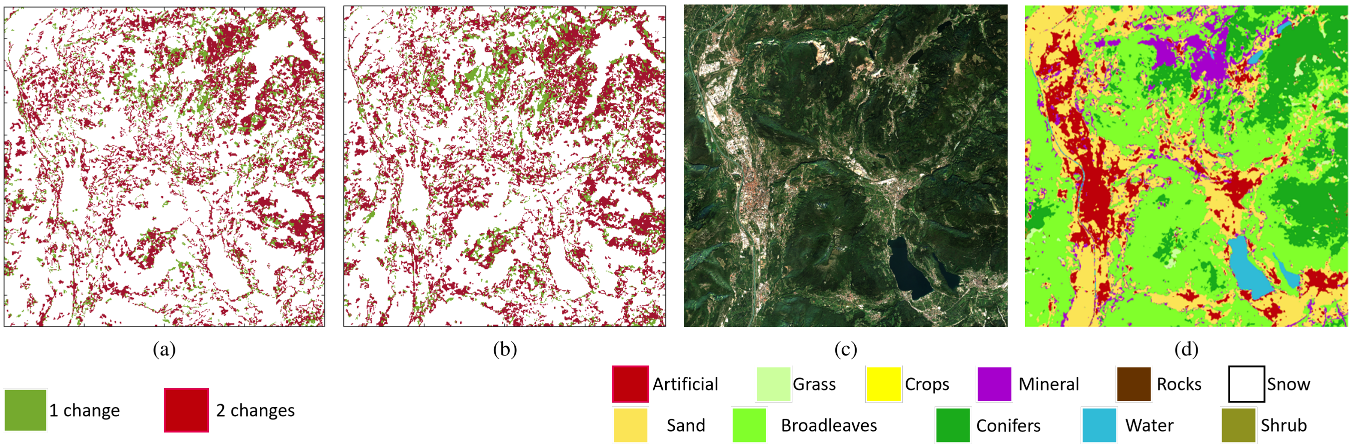

Fig. 2: Portion of the considered study area: (a) number of LC changes per pixel obtained with the proposed multi-year method, (b) number of LC changes per pixel obtained with the single year baseline method, (c) the true color composition of a Sentinel 2 image acquired in 2018, and (d) corresponding LC map obtained.

Table 1: LC map agreement achieved per class and overall considering the standard baseline method, i.e., single-year LSTM and the proposed multi-year LSTM.

Class	**Multi-Year LSTM**		**Single-Year LSTM**	
	2018-2019	2019-2020	2018-2019	2019-2020
Artificial	0.89	0.68	0.93	0.62
Grass	0.69	0.60	0.61	0.42
Crops	0.76	0.81	0.65	0.85
Mineral	0.33	0.51	0.34	0.43
Rocks	0.43	0.46	0.45	0.43
Sand	0.14	0.31	0.05	0.16
Broadleaves	0.85	0.87	0.79	0.84
Conifers	0.81	0.78	0.79	0.77
Shrubland	0.64	0.66	0.63	0.62
Water	0.97	0.98	0.98	0.95
Overall	0.78	0.78	0.74	0.74

as well as to reduce the detection of false changes with respect to the baseline method.

5. CONCLUSION

This paper presents a novel approach for the production of a TS of consistent LC maps by taking advantage of the temporal correlation existing between TSs acquired in different years in the same study area. The method first extracts an annual training set per TS to generate a set of preliminary LC maps. Then, it exploits the unchanged areas to define a reliable and informative multi-year training set to train a unique LSTM model. The preliminary results obtained demonstrate the effectiveness of the proposed approach. As future development, We plan to delve into the analysis of training methods to make use of longer TSs. A possible approach that can be investigated is the stateful LSTM to retain the cell state among prediction of sequences from the same pixel. The adoption of models that rely on both temporal and spatial correlation should also be considered for the creation of more consistent output maps.

6. REFERENCES

[1] K. D. Ngo, A. M. Lechner, and T. T. Vu, "Land cover mapping of the mekong delta to support natural resource management with multi-temporal sentinel-1a synthetic aperture radar imagery," *Remote Sensing Applications: Society and Environment*, vol. 17, pp. 100272, 2020.

[2] R. Malinowski, S. Lewiński, M. Rybicki, E. Gromny, M. Jenerowicz, M. Krupiński, A. Nowakowski, C. Wojtkowski, M. Krupiński, E. Krätzschmar, et al., "Automated production of a land cover/use map of europe based on sentinel-2 imagery," *Remote Sensing*, vol. 12, no. 21, pp. 3523, 2020.

[3] J. Inglada, A. Vincent, M. Arias, B. Tardy, D. Morin, and I. Rodes, "Operational high resolution land cover map production at the country scale using satellite image time series," *Remote Sensing*, vol. 9, no. 1, 2017.

[4] J. Wang, Y. Zhao, C. Li, L. Yu, D. Liu, and P. Gong, "Mapping global land cover in 2001 and 2010 with spatial-temporal consistency at 250m resolution," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 103, pp. 38–47, 2015, Global Land Cover Mapping and Monitoring.

[5] S. P. Abercrombie and M. A. Friedl, "Improving the consistency of multitemporal land cover maps using a hidden markov model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 2, pp. 703–713, 2015.

[6] S. M. Vicente-Serrano, F. Pérez-Cabello, and T. Lasanta, "Assessment of radiometric correction techniques in analyzing vegetation variability and change using time series of landsat images," *Remote Sensing of Environment*, vol. 112, no. 10, pp. 3916–3934, 2008.

[7] C. Paris and L. Bruzzone, "A novel approach to the unsupervised extraction of reliable training samples from thematic products," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 3, pp. 1930–1948, 2020.

[8] M. Bossard, J. Feranec, and J. Otahel, "CORINE land cover technical guide – Addendum 2000," Tech. Rep. 40, European Environment Agency, Copenhagen, 2000.