



HAL
open science

Unsupervised Learning of Robust Representations for Change Detection on Sentinel-2 Earth Observation Images

Michelle Aubrun, Andres Troya-Galvis, Mohanad Albughdadi, Romain Hugues, Marc Spigai

► **To cite this version:**

Michelle Aubrun, Andres Troya-Galvis, Mohanad Albughdadi, Romain Hugues, Marc Spigai. Unsupervised Learning of Robust Representations for Change Detection on Sentinel-2 Earth Observation Images. 13th International Symposium on Environmental Software Systems (ISESS), Feb 2020, Wageningen, Netherlands. pp.1-6, 10.1007/978-3-030-39815-6_1 . hal-03361874

HAL Id: hal-03361874

<https://inria.hal.science/hal-03361874>

Submitted on 1 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Unsupervised learning of robust representations for change detection on Sentinel-2 Earth Observation images

Aubrun Michelle¹, Troya-Galvis Andres¹, Albughdadi Mohanad², Hugues Romain¹
and Spigai Marc¹

¹Thales Alenia Space: 26 avenue Jean François Champollion, 31100 Toulouse, France
{michelle.aubrun, andres.troya-galvis, romain.hugues,
marc.spigai}@thalesaleniaspace.com

²TerraNIS: 12 avenue de l'Europe, 31520 Ramonville-Saint-Agne, France
mohanad.albughdadi@terrannis.fr

Abstract. The recent popularity of artificial intelligence techniques and the wealth of free and open access Copernicus data have led to the development of new data analytics applications in the Earth Observation domain. Among them, is the detection of changes on image time series, and in particular, the estimation of levels and superficies of changes. In this paper, we propose an unsupervised framework to detect generic but relevant and reliable changes using pairs of Sentinel-2 images. To illustrate this method, we will present a scenario focusing on the detection of changes in vineyards due to natural hazards such as frost and hail.

Keywords: Change detection, Unsupervised method, Artificial Intelligence, Sentinel-2 data, vineyard use case.

1 Introduction

With the advent of the Copernicus program and its wealth of free and open data, the Earth Observation (EO) domain is increasingly adopting automatic or semi-automatic data analytics applications, based on artificial intelligence techniques.

Using the spectral richness (13 bands), the fine temporal (few days) and spatial resolution (10 meters per pixel) of Sentinel-2, a lot of use cases can be carried out in di-verse sectors, and in particular those related to vineyard health assessment. Wine-making is one of the largest industries that represents a turnover of several billions of euros in France. However, this industry faces severe meteorological challenges such as frost and hail, that can cause significant loss in wine production. After such meteorological events, winemaking farmers need to evaluate the level of damage that occurred in their vineyards in order to receive subsidies from the state and the European insurances. Additionally, insurance companies must also estimate the damage levels using field visits in order to check the information provided, which is non-trivial and requires a huge budget and workforce. Coupling these Copernicus data with an appropriate change detection model will allow farmers and insurance companies to easily build a complete damage profile of the vineyards in case of any natural hazard.

In this context, a generic change detection application on Sentinel-2 time series was developed as part of CANDELA [1], an H2020 research and innovation project. Since the tool provides generic changes, many use cases can take full advantage of it in order to save time, effort and budget.

The rest of this article is structured as follows. Section 2 presents a brief review of existing change detection methods. Section 3 describes the methodology, and Section 4 is dedicated to the experiments and results obtained on the selected use case. Finally, Section 5 discusses our conclusions and future work perspectives.

2 Related work

Change detection is a well-known problem in the remote sensing community, thus several approaches have been developed to tackle this problem. The most classical approaches rely on pixel difference, image regression, image rationing, radiometric index difference, and metrics based on mutual information or correlation indexes [2]. Other approaches compute a distance metric on manually engineered feature spaces such as the Laws filters presented in [3]. Although these methods are fast and easy to interpret, they are sensitive to noise and often detect subtle but irrelevant changes.

Other approaches rely on machine learning to classify the image pixels into relevant classes and then verify that the class of a given area has not changed between the two images [4, 5]. Such methods provide additional information about the nature of the changes, but they are very task-specific and require labeled data in order to learn the classifier in the first place.

More recently, some approaches start making use of deep learning techniques such as Generative Adversarial Networks [6] or U-Net [7] to infer the change map directly from the two images.

Since labeled data is scarce and expensive to produce, we propose an unsupervised framework to detect generic but relevant changes between pairs of Sentinel-2 images.

3 Methodology

3.1 Change Detection Service

The proposed Change detection service has been implemented on Candela platform. As seen in Fig. 1, the pipeline is composed of several modules. The first one, named Jpeg2Tiff, consists in extracting the bands of interest, resampling them to get the same spatial resolution for all bands and concatenating them to obtain a single geoTiff image for each Sentinel-2 product. The second module, named Preprocess, verifies that all images provided represent exactly the same area and sorts them in the chronological order. And the last module, named ChangeDetect, computes the change detection maps for each pair of consecutive images of the time-series.

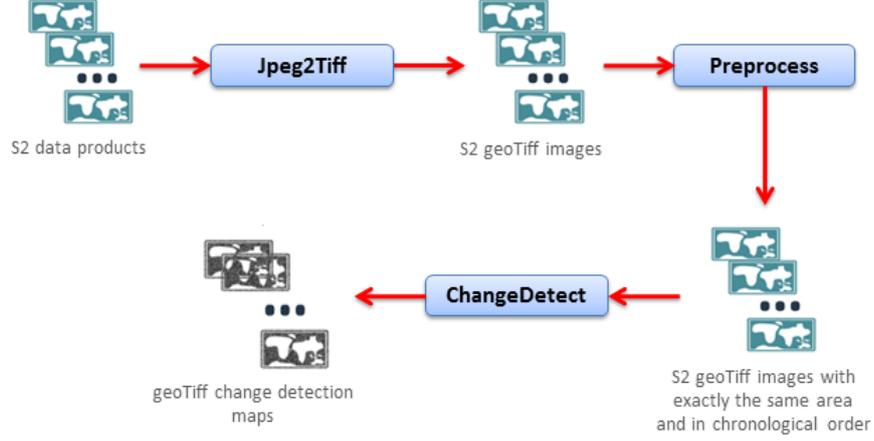


Fig. 1. Change detection pipeline

3.2 Framework

The proposed framework to generate the change detection map consists in projecting the image pixel space into a more robust feature space which is learnt by a neural network E . Then, computing a distance metric M between two encoded images in order to determine a change score. This whole framework can be described as follows:

$$C(X_1, X_2) = M(E(X_1), E(X_2)) \quad (1)$$

where X_1 and X_2 are the two images to be compared, which are split into individual $n \times n$ patches on which the encoding and the subsequent distance are computed.

3.3 Approach and implementation

In order to generate models of robust representations, an unsupervised approach based on stacked autoencoders [8, 9 & 10] was explored. Stacked autoencoders consist in two sub-networks, the first one is used to compress the input data into a fixed-length representation, while the second one tries to decompress this representation in order to obtain the initial data back. The networks are trained from only one image by minimizing the following loss:

$$L(x, e, d) = MSE(x, d(e(x))) \quad (2)$$

where x is an input image, e is the encoder network, d is the decoder network, and MSE is the mean squared error. In conclusion, the decoder network learns to reconstruct the input image from a compressed representation, forcing the encoder network to capture the main features of the image and discarding the noise.

Two different architectures of stacked autoencoders have been implemented:

- **Dense autoencoder** whose architecture makes use of one fully connected layer during the encoding and decoding steps.
- **Convolutional autoencoder** whose architecture is a fully convolutional network made of successive 3x3 convolution layers with 256 filters and stride 2 until the feature map size is reduced to 1x1. The decoder makes use of inverse convolutions until the feature map size is the same as the input image.

For both architectures, we used Adam for the optimizer function, a value of 0.001 for the learning rate, and a linear activation function for the last layer.

4 Experiments

4.1 Data description

To demonstrate the applicability of our approaches, a region of interest (ROI) near Bordeaux in France, well-known for its wine and affected by frost on 27th April 2017, was selected ($-0.3868367W$, $44.5202483N$: $0.1090724E$, $44.7963392N$). Two optical Sentinel-2 Level-2A (atmospherically corrected) data products of the T30TYQ tile with low cloud cover on the ROI were used for the analysis. One acquired during 19th April 2017 (before the frost) and the other during 29th April 2017 (after the frost). These products contain 13 spectral bands in the visible, near infrared and shortwave infrared part of the spectrum with different spatial resolutions.

To quantify the level of change (low, middle or high) at vineyard parcel level, a vector dataset containing 11355 parcels from the French Parcel Registration System [11] was also used.

4.2 Settings

The training data come from a Sentinel-2 Level-2A image from Toulouse with low cloud cover. This image have been tiled into normalized patches of $5 \times 5 \times N$ pixels with N the number of bands. For our vineyard use case, we have decided to consider all the bands at 10 and 20 meters of resolution, what represents 10 bands, and resample them according to the blue band. Thus, our models were trained on 65536 patches of $5 \times 5 \times 10$ pixels selected randomly by the algorithm for 10 epochs. The models are based on the architectures presented in section 3.3 and provide a 25-dimensional vector that corresponds to an encoded representation of the patch in another feature space. Thus, the shapes of the different layers for the dense encoder are $5 \times 5 \times 10 \rightarrow 1 \times 250 \rightarrow 1 \times 25$, and for the convolutional encoder are $5 \times 5 \times 10 \rightarrow 3 \times 3 \times 256 \rightarrow 2 \times 2 \times 256 \rightarrow 1 \times 1 \times 256 \rightarrow 1 \times 1 \times 25$. Both models are available on Candela platform.

The testing data correspond to the two images of our vineyard use case. The same pre-processing procedure as the training image was applied to encode all possible patches. Finally, the L2 distance between each pair of encoded feature vectors have been chosen to measure the amount of change.

4.3 Results

To evaluate the performance of our method, we have asked an expert, who has a dual competence in remote sensing and agronomy and whose job is to extract accurate and relevant information from images, to create a ground truth of change levels from the Sentinel-2 Level-2A images. As the analysis takes a lot of time, the ground truth has been done on an area that contains 253 parcels. Fig. 2 shows the ground truth, the results generated by our framework with using both models described in section 4.2 and the result generated by our framework without applying any model. We can see that the results generated by our approach seem more similar to the ground truth than the approach where data are not projected into a learnt feature space. To quantify these similarities, we have extracted the median values of the different change detection maps at parcel level and considered the correlation coefficients as the evaluation metric. The correlation coefficient values were 0.79, 0.80 and 0.56 for respectively the frameworks with the dense encoder, the convolutional encoder and without encoder. These scores prove the efficiency of our approaches compared with classic approaches.

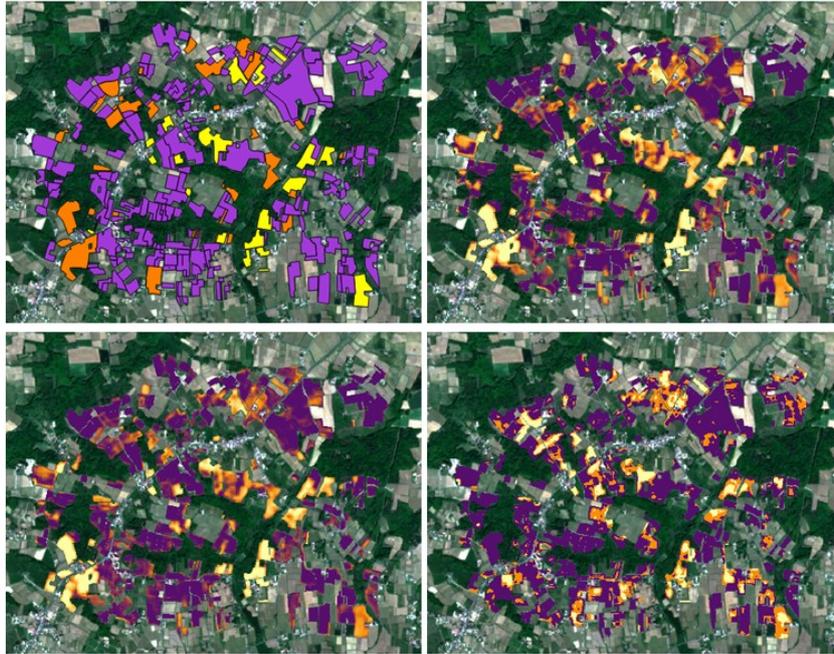


Fig. 2. Comparison of change detection results get by a Remote Sensing Expert (up-left), dense encoder (up-right), convolutional encoder (down-left) and without encoder (down-right). Change levels; violet: low, orange: middle, yellow: high

5 Conclusion

In this study, we have demonstrated on a real use case the effectiveness of our unsupervised approaches to provide trustworthy change detection maps. These approaches lead to generic change level maps, but coupling with vector datasets, the type of changes can be specified and the percentage of changes estimated. Thus, these results may facilitate the work of many operators in different sectors.

Acknowledgements. The work presented in this paper was supported by the H2020 CANDELA project under grant agreement No.776193

References

1. CANDELA project, <http://www.candelah2020.eu/>
2. Ashbindu, S.: Review Article Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing* 10(6), 989-1003 (1989)
3. Kenneth, L.: Textured Image Segmentation, Ph.D. Dissertation, University of Southern California (1980)
4. Serra, P., Pons, X., Sauri, D.: Post-classification change detection with data from different sensors: Some accuracy considerations. *International Journal of Remote Sensing* 24(16), 3311-3340 (2003)
5. Alboody, A., Sedes, F., Inglada, J.: Post-classification and spatial reasoning: new approach to change detection for updating GIS database. In: 3rd International Conference on Information and Communication Technologies, pp. 1-7. From Theory to Applications, Damascus (2008)
6. Maoguo, G., Xudong, N., Puzhao, Z., Zhetao, L.: Generative Adversarial Networks for Change Detection in Multispectral Imagery. *IEEE geoscience and remote sensing letters* 14(12), 2310-2314 (2017)
7. Louis de Jong, K., Sergeevna Bosman, A.: Unsupervised Change Detection in Satellite Images Using Convolutional Neural Networks. *International Joint Conference on Neural Networks*. Budapest (2019)
8. LeCun, Y.: Modèles connexionistes de l'apprentissage. PhD thesis, Université de Paris VI (1987)
9. Planinsic, P., Gleich, D.: Temporal Change Detection in SAR Images Using Log Cumulants and Stacked Autoencoder. *IEEE geoscience and remote sensing letters* 15(2), 1-5 (2018)
10. Mou, L., Ghamisi, P., Xiang Zhu, X.: Unsupervised Spectral-Spatial Feature Learning via Deep Residual Conv-Deconv Network for Hyperspectral Image Classification. *IEEE transactions on geoscience and remote sensing* 56(1), 391-406 (2018)
11. French Parcel Registration System, <https://www.data.gouv.fr/en/datasets/registre-parcellaire-graphique-rpg-contours-des-parcelles-et-ilots-cultureaux-et-leur-groupe-de-cultures-majoritaire/>