SPATIOTEMPORAL SEGMENTATION OF SATELLITE IMAGE TIME SERIES USING SELF-ORGANIZING MAP

Baggio Luiz de C. e Silva[1]*, Felipe C. de Souza[1], Karine R. Ferreira[1], Gilberto R. de Queiroz[1], Lorena A. dos Santos[1]

[1] National Institute for Space Research (INPE)
Avenida dos Astronaustas, 1758, São José dos Campos, SP, Brazil.
(baggio.silva, felipe.carvalho, karine.ferreira, gilberto.queiroz, lorena.santos)@inpe.br

KEY WORDS: Spatiotemporal Segmentation, Satellite Image Time Series, Land Use and Land Cover Changes, Unsupervised Classification, Clustering Algorithms, Earth Data Cubes.

ABSTRACT:

Nowadays, researchers have free access to an unprecedentedly large amount of remote sensing images collected by satellites and sensors with different spatial, temporal, and spectral resolutions. This scenario has promoted the use of satellite image time series for spatiotemporal analysis. This paper presents a methodology for spatiotemporal segmentation of satellite image time series. Spatiotemporal segmentation finds homogeneous regions in space and time from remote sensing images based on spectral features. The proposed approach is unsupervised based on the self-organizing map (SOM) neural network and hierarchical clustering algorithm. It was implemented and applied to a region in the Mato Grosso state, Brazil. The results were evaluated using qualitative and quantitative approaches. In the qualitative approach, visual analysis was performed based on the land use and land cover map of the TerrraClass Cerrado project. In the quantitative approach, supervised and geometric metrics were used to analyze the quality of the produced segments. The results obtained are promising since the segments produced were homogeneous and with a low occurrence of over-segmentation.

1. INTRODUCTION

Remote sensing imagery is essential for monitoring environmental changes, natural disasters, population growth, and many other applications. Nowadays, a large number of Earth observation satellites have generated big volumes of images with different spatial and temporal resolutions every day. In 2019, the volume of open data produced by Landsat-7 and Landsat-8, MODIS (Terra and Aqua units), and the three first Sentinel satellites (Sentinel-1, -2 and -3) were around 5 PB(Soille et al., 2018). This scenario poses a great challenge for storing, processing, and analyzing this large amount of data and also promotes the use of satellite image time series for spatiotemporal analysis(Simoes et al., 2021a).

Spatiotemporal segmentation consists in finding homogeneous regions in space and time from remote sensing images based on spectral features (Petitjean et al., 2012). Remote sensing image segmentation is widely used in object-based approaches to produce homogeneous geographical objects (Xi et al., 2019). Despite being an important step in remote sensing image processing, there are few methods proposed in the literature that consider the space and time dimensions to produce homogeneous segments (Petitjean et al., 2012, Costa et al., 2018b, Xi et al., 2019, Fare Garnot and Landrieu, 2021)

Petitjean et al. (2012) propose an approach that segments each image of the series and then combines these segments to provide a classification map. Costa et al. (2018) propose a method for segmentation applied to image time series, adapting the traditional region growing algorithm and using the Dynamic Time Warping (DTW) distance to detect homogeneous regions in space and time. Xi et al. (2019) describe a multiresolution segmentation to generate multiscale spatiotemporal cubes. Garnot and Landrieu (2021) present a method that processes a sequence of images in parallel by a shared convolutional encoder. At the lowest resolution, a temporal encoder produces a set of temporal attention masks for each pixel, then spatially interpolated at all resolutions.

Different from the existing methods, this paper presents a spatiotemporal segmentation of satellite image time series using a self-organizing map (SOM) and hierarchical clustering algorithm. SOM is an unsupervised artificial neural network based on competitive learning that reduces a high dimensional feature input space onto a lower-dimensional feature output space (Kohonen, 1998). It also preserves the neighborhood topology; that is, similar input data are mapped to the same neuron or a nearby one. SOM has been widely used for image time series clustering and spatiotemporal patterns discovering (Zurita-Milla et al., 2013) and (Santos et al., 2021).

The proposed approach is unsupervised and applied to satellite image time series; that is, time series associated with pixels of a satellite image sequence ordered in time. Based on a given threshold or a number of distinct clusters, it computes the similarity among all image time series to produce homogeneous regions in space and time. The proposed method was implemented and tested in a region in the Mato Grosso state, Brazil, producing good results that are described in this paper. The code and data sets used in this work are available on Github to promote reproducibility[1].

An approach to finding homogeneous regions in space and time using satellite image time series is very useful for identifying some types of land use and cover classes. Some classes, such as one-cycle and two-cycle agriculture types, are correctly identified when we consider the temporal variation of satellite image spectral values. Regions that are used for one-cycle and two-cycle crops are homogeneous during an agriculture year. Thus,

* Corresponding author

[1] https://github.com/brazil-data-cube/st-segmentation

it is important to consider the spatial and temporal dimensions to extract these classes from satellite image sequences ordered in time.

This paper is organized as follows: Section 2 presents the clustering methods used in spatiotemporal segmentation. Section 3 describes the methodology with the steps of clustering tasks and spatiotemporal segmentation. Section 4 explores a qualitative and quantitative evaluation of the segments and highlights the main results. Section 6 describes the settings used for the experiments, run times, and final considerations.

2. CLUSTERING METHODS

The method is based on two unsupervised clustering algorithms: SOM and agglomerative hierarchical clustering. The following describes each method in detail.

2.1 Self-Organizing Map (SOM)

SOM is an unsupervised learning method based on competitive learning that reduces a high-dimensional feature input space to a low-dimensional feature output space. A dataset can be mapped and represented by a set of neurons by using weight vectors. Preserve neighborhood topology is a key feature of SOM so that similar input data is mapped to the same or a neighboring neuron.

Each output space neuron j has a weight vector (codebook vector) $w_j = [w_{j1}, \ldots, w_{jn}]$ that contains the same dimension n of the input data $x(t) = [x(t)_1, \ldots, x(t)_n]$. The algorithm consists of two main steps. First, the distances D_j between a sample and each neuron in the SOM grid are calculated. The neuron with the smallest distance value d_b is selected as the best matching unit (BMU) for that sample. The equations for calculating distance and BMU are

$$D_j = \sqrt{\sum_{i=1}^{n}(x(t) - w_{ji})^2} \tag{1}$$

$$d_b = min(D_1, ..., D_j) \tag{2}$$

The second step is to adjust the weight vectors of the BMU and its neighbors so that the neurons have similar properties. The equation for the adjustment is

$$w_{ji} = w_{ji} + \alpha \times h_{b,j}[x(t)_i - w_{ji}]. \tag{3}$$

Where α is the learning rate and $h_{b,j}$ is the neighborhood function.

These steps are repeated T times to ensure that the neurons organize themselves into a similar neighborhood. Then, one neuron is assigned for each sample in the dataset (Kohonen, 1998).

2.2 Hierarchical Clustering

Hierarchical clustering is a method that partitions data sequentially, creating a hierarchy of clusters (Everitt et al., 2011). This representation makes it easier to visualize each step where level similarity occurs. Hierarchical clustering is classified into two types: agglomerative and divisive. The divisive algorithm begins with the entire dataset in one cluster and then divides it into two more similar clusters. The agglomerative method starts with each weight vector in its own cluster, computes a similarity matrix, and identifies the two most similar groups. The clusters are merged at each step, and the hierarchy is built using linkage criteria (Leonard Kaufman, 1990).

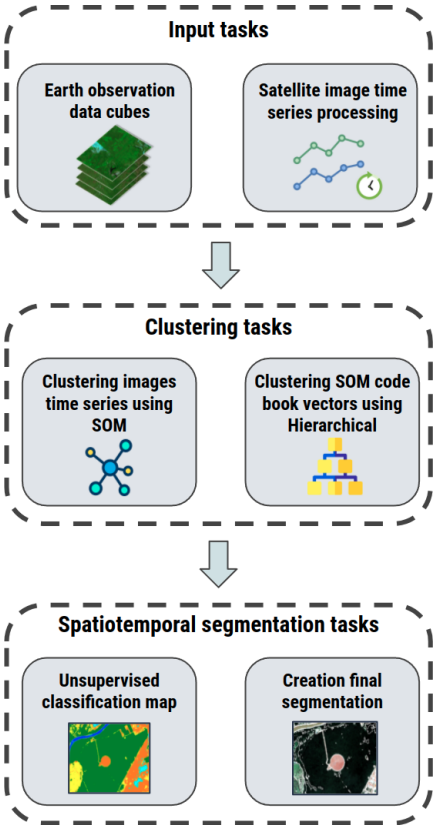

Figure 1. Methodology for spatiotemporal segmentation of satellite image time series.

Hierarchical algorithms construct a binary tree called a dendrogram. This structure represents the order in which the clusters were merged. A dendrogram divides the data into groups that are homogeneous within themselves. The variability of the data can be visualized using the tree hierarchy. The objective of the dendrogram is to explore and define the appropriate number of clusters based on the level of analysis.

3. MATERIAL AND METHODS

Figure 1 shows the methodology for spatiotemporal segmentation using SOM and agglomerative hierarchical clustering. The input data are a time series of satellite images obtained over a data cube, consisting of one or more spectral bands. After applying the SOM method, the algorithm's output provides weight vectors, also called codebook vectors, which have the same size as the input time series. The number of codebook vectors depends on the grid size defined for the SOM, where each vector represents a set of time series. In the agglomerative hierarchical clustering step, the codebook vectors are the input data in which the data are clustered into levels based on a similarity score. After the final number of clusters is determined, the unsupervised map is created in the spatiotemporal segmentation tasks. However, smoothing is performed on this map, which is used as the basis for segmentation.

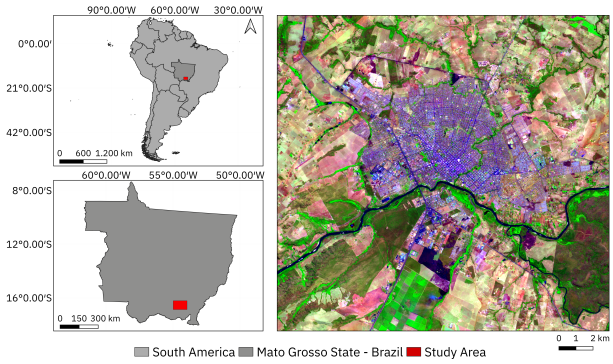

Figure 2. The study area where the experiments were conducted. The image on the right was taken from Sentinel-2/MSI with a 10-meter resolution. The agricultural composition (Swir16, Nir, and Blue) was used.

3.1 Study Area

The study area is in the Mato Grosso state, located in the mid-west of Brazil. Mato Grosso is one of the largest states in Brazil; its territory is composed of three biomes: Cerrado, Amazon, and Pantanal. Therefore, it presents a high spatial and temporal heterogeneity. Figure 2 shows the study area that is located in Rondonópolis, a city in the Mato Grosso state. The image on the right side corresponds to the satellite/sensor Sentinel-2/MSI from an agricultural composition (Swir16, Nir, and Blue) with a 10-meter spatial resolution.

The selected study region has several interesting land use and cover elements for the segmentation task. For example, an extensive urban center, the watershed of the Rio Vermelho that extends throughout the state, native vegetation, agricultural areas, and exposed soil regions.

3.2 Earth observation data cubes

Earth observation (EO) data cubes can be defined as multidimensional arrays with space and time dimensions and spectral derived properties created from remote sensing images mainly to support image time series analysis (Appel and Pebesma, 2019). The Brazil Data Cube (BDC) project is producing EO data cubes from analysis-ready data of CBERS-4, Sentinel-2, and Landsat-8 satellite images for the entire Brazilian territory (Ferreira et al., 2020).

This paper used the EO data cubes of Sentinel-2 images with a spatial resolution of 10-meters and temporal aggregation of 16 days using the best-pixel approach produced by the BDC project. For the experiments, we used these EO data cubes for the study area considering one agricultural year, 2019-08-29 to 2020-08-12, corresponding to 23 images.

To be general as possible, only the Normalized Difference Vegetation Index (NDVI) was used for this study. However, if the goal is to segment a specific land cover, different bands can be used to segment the target better.

3.3 Satellite image time series processing

The step of satellite image time series processing decreases the cloud and cloud shadow in the image time series. The pixels flagged as clouds or cloud shadow by the cloud band were marked as invalid (NAS). For each non-valid pixel, a linear interpolation was applied, in which the previous or later times of the same pixel were considered to fill the non-valid values. Figure 3 shows a time series of NDVI extracted from the EO data cube of Sentinel-2 images and associated with a spatial location of the study area.

After interpolation, the image time series were formatted to serve as input to the SOM. Based on the study area section, the images used in this work have 1954×2188 pixels. The input format of the model is an array$[4275352, 23]$, where the first component is the total size image's pixels and the second one depends on the temporal resolution and spectral bands. For this experiment, 23 time-series images with only one band were used. Another configuration, for example, if we had a temporal resolution of 25 images and three spectral bands, would be the array $[4275352, 75]$, since multiplying 25 by 3 gives 75, which corresponds to the concatenation of spectral bands.

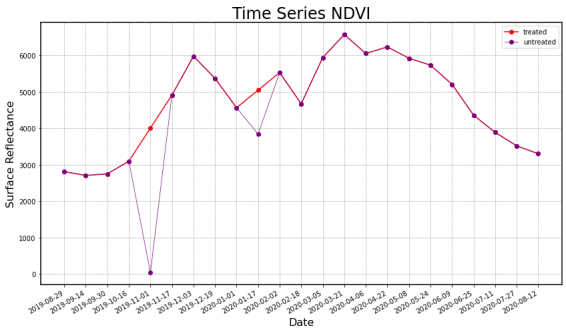

Figure 3. Example of an interpolated NDVI time series.

3.4 Clustering and Segmentation

Each time series represents one image pixel in the time domain and serves as input for the SOM algorithm. After the clustering step, each time series is assigned to a neuron, which in turn has an associated codebook vector, i.e., the codebook vector represents all-time series associated with the respective neuron, that characterizes the dimensionality reduction since an image with N pixels is represented by the number of neurons that make up the model.

Note that the number of patterns found in the image depends on the grid size chosen for SOM. How they are grouped also depends on the selected distance, the shape of the grid, and the neighborhood function.

After clustering with the SOM, the codebook vectors resulting from the first step are used as input for agglomerative hierarchical clustering. This is a clustering of the entire image represented by the codebook vectors, which would be impractical if all pixels were used, depending on the size of the image. After the new clustering, it is possible to create a visualization of the different levels of the clusters formed using a dendrogram. As in the previous step, the clusters are formed depending on some parameters such as the distance to be computed, the linkage criterion, and the number of the new clusters.

Once the unsupervised map was created using the clustered hierarchical method, spatial smoothing was applied to reduce noise. Since this is a pixel-based approach, one pixel or a small set of noisy pixels may interfere with the formation of polygons more representative of the actual area. Finally, the last step is the generation of homogeneous segments to their temporal profile.

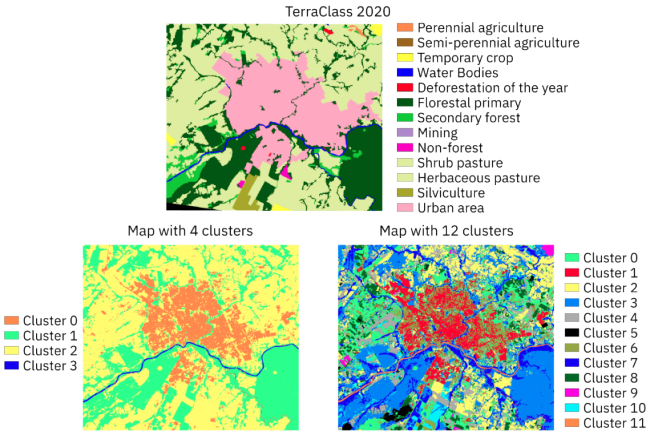

Figure 4. Results obtained with the settings with 4 and 12 groups.

Figure 5. Spatiotemporal segmentation results (red outlines).

4. RESULTS AND DISCUSSION

For this experiment with the following results, only two parameters were selected for the SOM, the grid size and the number of epochs, and two parameters were also selected for the hierarchical algorithm, the final number of clusters and the linkage, the other parameters were set by default for `XPySom` (Mancini et al., 2020) and `sklearn` (Pedregosa et al., 2011) Agglomerative Clustering, respectively. For the SOM with 20 epochs, a grid with 12x12 neurons was chosen, and for the hierarchical algorithm, two configurations with 4 and 12 final cluster numbers with the average linkage were used.

The results were evaluated using qualitative and quantitative approaches. In the qualitative approach, a visual analysis of the satellite image of a sentinel-2 image and a comparison with the reference provided by the TerraClass Cerrado project for the same period. In the quantitative approach, supervised and geometric metrics were used to analyze the quality of the produced segments. The following describes the qualitative and quantitative results, respectively.

4.1 Visual analysis

At the first moment, a qualitative assessment compares the segments with a sentinel-2 image and the TerraClass Cerrado map based on visual interpretation. TerraClass Cerrado project produces land use and land cover maps for the Cerrado biome in Brazil with spatial resolution between 20 and 30 meters (Almeida et al., 2016). This project's methodology for producing the land use and land cover map is based on several steps, mainly visual analysis by remote sensing experts. The experts use MODIS time series and high-resolution imagery to distinguish the different agricultural cycles in each Brazilian biome. These data sets are available through the TerraClass portal[2].

The unsupervised result for the entire scene is shown in Figure 4. In this Figure, we present the thematic map of the TerraClass project for the 2020 year and the unsupervised maps with 4 and 12 clusters. Note that each map has its legend, and the colors of the unsupervised maps do not correspond to the TerraClass map classes.

In general, the two configurations were able to capture the main features of the region, namely: urban area, native vegetation, and water bodies. In the configuration with 4 groups, it was possible to identify urban areas and waters more homogeneously. However, regions with different vegetation types, cropland, and pasture land obtained a greater integration but a good separation from the other classes.

To support visual analysis and understand the dynamics of the model with 4 groups, Figure 7 shows the dendrogram created by the hierarchical clustering method, where 144 (12x12 SOM grid) codebooks vectors were integrated into 4 final clusters defined only by the time series distance. Figure 8 shows the NDVI profiles for each codebook vector with the respective colors assigned to each final cluster. Furthermore, Figure 5 presents the segments generated from this map. It is possible to notice how the rivers and regions of dense vegetation are well delineated.

In the model with 12 groups, it was observed that it was possible to detect more nuances in land use and land cover classes. For example, the model captured different types of roofs, containers, and roads in the urban region. Another example is the detected vegetation types, where it was possible to distinguish between dense and shallow vegetation.

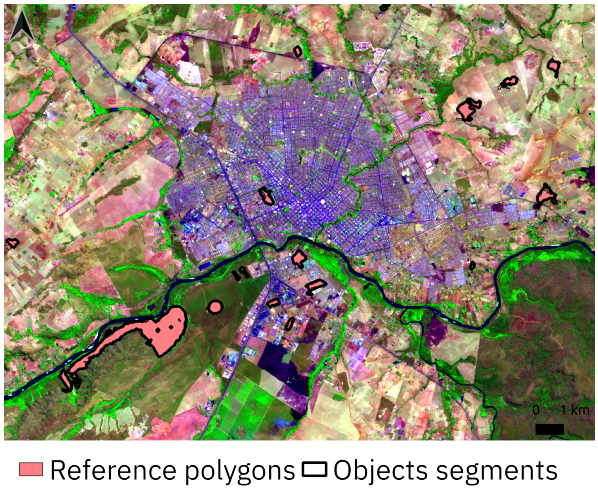

Figure 6. Reference Polygons in red used for quantitative evaluation with a sentinel-2 agricultural composition.

4.2 Supervised accuracy assessment

In remote sensing image segmentation applications, determining the accuracy of segmentation undertaken quantitatively is

[2] https://www.terraclass.gov.br/

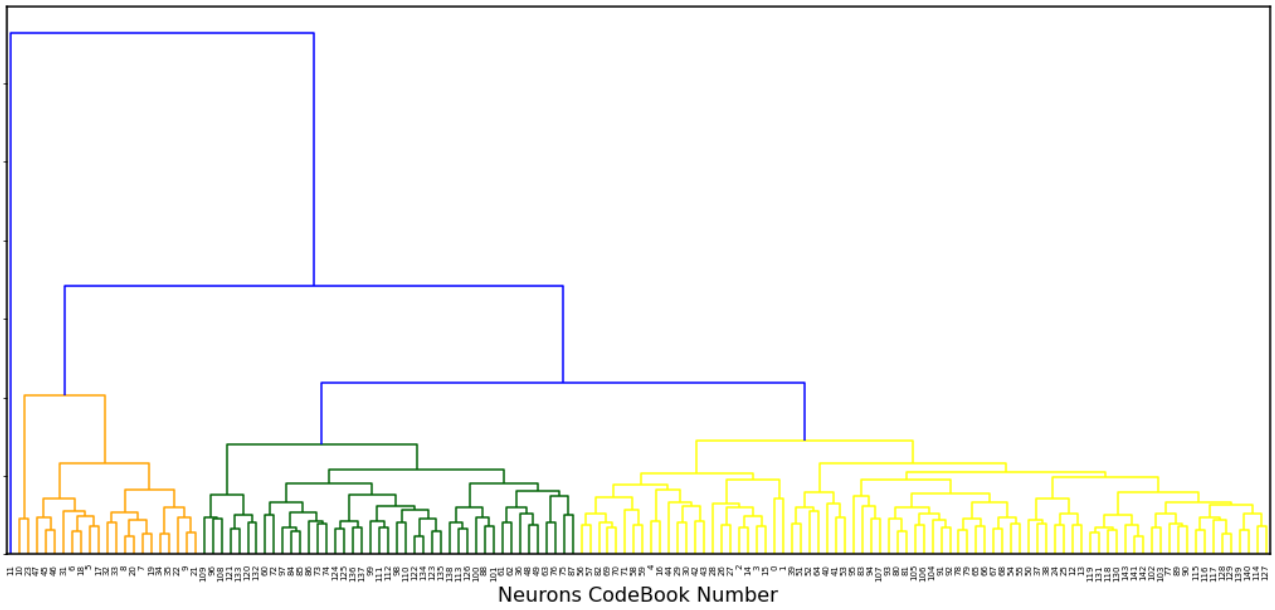

Figure 7. Dendrogram Results.

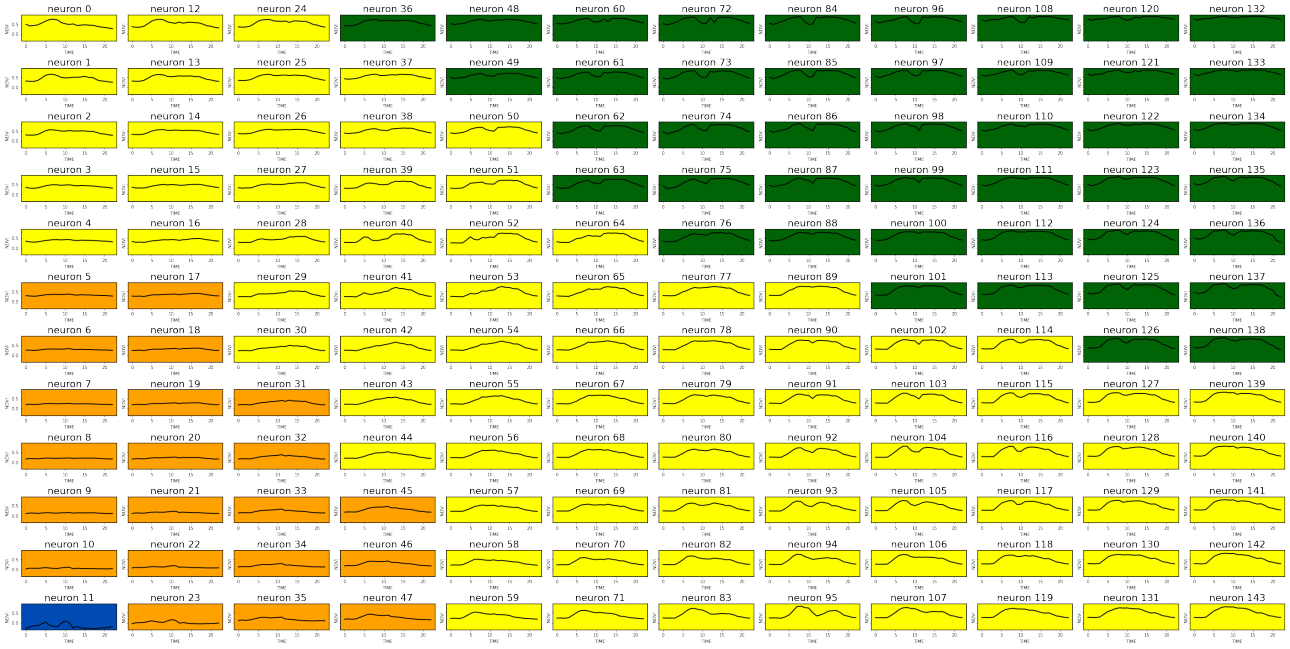

Figure 8. NDVI Codebooks vector of SOM.

still at an early stage of maturation (Costa et al., 2018a). Since much of the evaluated work uses a qualitative approach to measure the quality of its results, this paper used supervised metrics for quantitative analysis to determine segmentation accuracy to perform the quantitative analysis. Supervised metrics measure the similarity of generated segments to reference polygons by shape similarity. These metrics are divided into area-based and location-based. The area-based metrics evaluate the overlap region between the segmentation object and the reference polygon. The location-based metrics evaluate the distance between the centroids of the segmentation objects and the reference polygon (Clinton et al., 2010).

Costa et al. (2018a) recommends the use of area-based metrics for applications in which geometric representation is more important than the thematic information associated with segment objects. Therefore, in this paper, we used the metrics Precision (van Rijsbergen, 1979, Zhang, 1996), Recall (van Rijsbergen, 1979, Zhang, 1996), F_measure (van Rijsbergen, 1979, Zhang, 1996), and Match (M) (Janssen and Molenaar, 1995, Feitosa et al., 2010). The Precision metric measures the precision of the segment object, with the reference polygon being sensitive to under-segmentation. The Recall metric measures how similar the reference polygon is to the segment object and is more sensitive to over-segmentation. The M metric measures the average match between polygons and segments, measuring under- and over-segmentation. Finally, F_measure is a combined metric representing the balance between recall and precision, which also measures under- and over-segmentation. These metrics range from 0 and 1, with the optimal value being 1.

The vegetation polygons were extracted based on the Terra-

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-V-3-2022-255-2022 | © Author(s) 2022. CC BY 4.0 License.

Class thematic map and compared with the 4 clusters map to perform the quantitative assessments. Figure 6 shows the reference polygons extracted from the TerraClass Cerrado project and used as reference polygons. The segmentation results compared to the reference polygons can be seen in the same Figure, where the red area represents the reference polygons, and the white lines represent the segment's object. To calculate the metrics, we used the R package `segmetric` (Simoes et al., 2021b).

Metric	Value
Recall	0.87
Precision	0.83
F_measure	0.85
M	0.64

Table 1. Quantitative results of the spatiotemporal segmentation.

The results of the metrics evaluated to determine the accuracy of the generated segmentation are shown in Table 1. It can be noted that the Recall metric presented the best result (0.87), indicating a small over-segmentation error. The M metric obtained the worst result since this metric calculates the average between the match of the two geometries. This value is due to some shapes that obtained a low value. The results obtained were satisfactory and showed a good agreement with the reference polygons.

5. FINAL COMMENTS

This paper presents a spatiotemporal segmentation based on two clustering algorithms. Two configurations were used to present the model, and qualitative and quantitative evaluations were made to analyze the segments generated by the algorithm.

Note that due to the nature of spatiotemporal segmentation, some polygons represent land use and land cover classes that are often difficult to evaluate with a single image. On the other hand, land use and land cover types that have characteristic temporal profiles, such as agriculture, deforestation, and forest fires, can be represented by a segment. Thus, with the knowledge of experts, it is possible to generate segments of semantic information analyzing these temporal profiles.

The model was run on a Google Collab virtual machine with the following specifications: Intel(R) Xeon(R) CPU @ 2.20GHz with 2 cores, 12GB ram and 25GB disk space, Nvidia K80 / T4 GPU with 12GB/16GB GPU Memory, which obtained the times of 15 to 20 seconds for training with 20 epochs and approximately 5 seconds for prediction for the SOM algorithm which had the longest execution time. Further study will be done to take into account neighboring pixels for contextual information and explore different spectral bands, temporal resolutions, and other spatial resolutions for different land use and land cover applications.

6. ACKNOWLEDGEMENTS

This study was financed in part by the Coordenação de Aperfeçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and by the Environmental Monitoring of Brazilian Biomes project (Brazil Data Cube), funded by the Amazon Fund through the financial collaboration of the Brazilian Development Bank (BNDES) and the Foundation for Science, Technology and Space Applications (FUNCATE), Process 17.2.0536.1.

REFERENCES

Almeida, C. A., Coutinho, A. C., Esquerdo, J. C. d. M., Adami, M., Venturieri, A., Diniz, C. G., Dessay, N., Durieux, L., Gomes, A. R., 2016. High spatial resolution land use and land cover mapping of the Brazilian legal Amazon in 2008 using Landsat-5/TM and MODIS data. *Acta Amazonica*, 46(3), 291–302.

Appel, M., Pebesma, E., 2019. On-demand processing of data cubes from satellite image collections with the gdalcubes library. *Data*, 4(3).

Clinton, N., Holt, A., Scarborough, J., Yan, L. I., Gong, P., 2010. Accuracy assessment measures for object-based image segmentation goodness. *Photogrammetric Engineering and Remote Sensing*, 76(3), 289–299.

Costa, H., Foody, G. M., Boyd, D. S., 2018a. Supervised methods of image segmentation accuracy assessment in land cover mapping. *Remote Sensing of Environment*, 205(December 2016), 338–351. https://doi.org/10.1016/j.rse.2017.11.024.

Costa, W. S., Fonseca, L. M. G., Korting, T. S., Bendini, H. D. N., Cartaxo Modesto De Souza, R., 2018b. Spatiooral segmentation applied to optical remote sensing image time series. *IEEE Geoscience and Remote Sensing Letters*, 15(8), 1299–1303.

Everitt, B., Landau, S., Leese, M. S., 2011. *D. Cluster Analysis, 5th ed.; Wiley Series in Probability and Statistics*. Wiley: Hoboken, NJ, USA.

Fare Garnot, V. S., Landrieu, L., 2021. Panoptic segmentation of satellite image time series with convolutional temporal attention networks. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 4852–4861.

Feitosa, R., Ferreira, R., Almeida, C., Camargo, F., Costa, G., 2010. Similarity metrics for genetic adaptation of segmentation parameters. *Geographic Object-Based Image Analysis (GEOBIA) 2010 conference held*, 29.

Ferreira, K. R., Queiroz, G. R., Vinhas, L., Marujo, R. F., Simoes, R. E., Picoli, M. C., Camara, G., Cartaxo, R., Gomes, V. C., Santos, L. A., Sanchez, A. H., Arcanjo, J. S., Fronza, J. G., Noronha, C. A., Costa, R. W., Zaglia, M. C., Zioti, F., Korting, T. S., Soares, A. R., Chaves, M. E., Fonseca, L. M., 2020. Earth observation data cubes for Brazil: Requirements, methodology and products. *Remote Sensing*, 12(24), 1–19.

Janssen, L. L., Molenaar, M., 1995. Terrain Objects, Their Dynamics and Their Monitoring by the Integration of GIS and Remote Sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 33(3), 749–758.

Kohonen, T., 1998. The self-organizing map. *Neurocomputing*, 21(1-3), 1–6.

Leonard Kaufman, P., 1990. *Finding Groups in Data: An Introduction to Cluster Analysis, 9th ed.* Wiley: Hoboken, NJ, USA.

Mancini, R., Ritacco, A., Lanciano, G., Cucinotta, T., 2020. Xpysom: High-performance self-organizing maps. *2020 IEEE 32nd International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)*, 209–216.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Petitjean, F., Kurtz, C., Passat, N., Gançarski, P., 2012. Spatio-temporal reasoning for the classification of satellite image time series. *Pattern Recognition Letters*, 33(13), 1805–1815.

Santos, L. A., Ferreira, K., Picoli, M., Camara, G., Zurita-Milla, R., Augustijn, E.-W., 2021. Identifying Spatiotemporal Patterns in Land Use and Cover Samples from Satellite Image Time Series. *Remote Sensing*, 13(5). https://www.mdpi.com/2072-4292/13/5/974.

Simoes, R., Camara, G., Queiroz, G., Souza, F., Andrade, P. R., Santos, L., Carvalho, A., Ferreira, K., 2021a. Satellite Image Time Series Analysis for Big Earth Observation Data. *Remote Sensing*, 13(13). https://www.mdpi.com/2072-4292/13/13/2428.

Simoes, R., Sanchez, A., Picoli, M., 2021b. segmetric: Metrics for Assessing Segmentation Accuracy for Geospatial Data. INPE - Brazilian National Institute for Space Research, Sao Jose dos Campos, Brazil.

Soille, P., Burger, A., De Marchi, D., Kempeneers, P., Rodriguez, D., Syrris, V., Vasilev, V., 2018. A versatile data-intensive computing platform for information retrieval from big geospatial data. *Future Generation Computer Systems*, 81, 30–40. https://doi.org/10.1016/j.future.2017.11.007.

van Rijsbergen, C., 1979. Information retrieval, 2nd edbutterworths.

Xi, W., Du, S., Wang, Y. C., Zhang, X., 2019. A spatiotemporal cube model for analyzing satellite image time series: Application to land-cover mapping and change detection. *Remote Sensing of Environment*, 231(May), 111212. https://doi.org/10.1016/j.rse.2019.111212.

Zhang, Y. J., 1996. A survey on evaluation methods for image segmentation. *Pattern Recognition*, 29(8), 1335–1346.

Zurita-Milla, R., Van Gijsel, J. A., Hamm, N. A., Augustijn, P. W., Vrieling, A., 2013. Exploring spatiotemporal phenological patterns and trajectories using self-organizing maps. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4), 1914–1921.