



Article

Mapping Deforestation in Cerrado Based on Hybrid Deep Learning Architecture and Medium Spatial Resolution Satellite Time Series

Bruno Menini Matosak ^{1,*}, Leila Maria Garcia Fonseca ¹, Evandro Carrijo Taquary ¹, Raian Vargas Maretto ², Hugo do Nascimento Bendini ¹ and Marcos Adami ¹

¹ Earth Observation and Geoinformatics Division, National Institute for Space Research (INPE), Avenida dos Astronautas 1758, Jardim da Granja, Sao Jose dos Campos 12227-010, Brazil; leila.fonseca@inpe.br (L.M.G.F.); evandro.taquary@inpe.br (E.C.T.); hugo.bendini@inpe.br (H.d.N.B.); marcos.adami@inpe.br (M.A.)

² Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Hengelosestraat 99, 7514AE Enschede, The Netherlands; r.v.maretto@utwente.nl

* Correspondence: bruno.matosak@inpe.br

Abstract: Cerrado is the second largest biome in Brazil, covering about 2 million km². This biome has experienced land use and land cover changes at high rates due to agricultural expansion so that more than 50% of its natural vegetation has already been removed. Therefore, it is crucial to provide technology capable of controlling and monitoring the Cerrado vegetation suppression in order to undertake the environmental conservation policies. Within this context, this work aims to develop a new methodology to detect deforestation in Cerrado through the combination of two Deep Learning (DL) architectures, Long Short-Term Memory (LSTM) and U-Net, and using Landsat and Sentinel image time series. In our proposed method, the LSTM evaluates the time series in relation to the time axis to create a deforestation probability map, which is spatially analyzed by the U-Net algorithm alongside the terrain slope to produce final deforestation maps. The method was applied in two different study areas, which better represent the main deforestation patterns present in Cerrado. The resultant deforestation maps based on cost-free Sentinel-2 images achieved high accuracy metrics, peaking at an overall accuracy of $99.81\% \pm 0.21$ and F1-Score of 0.8795 ± 0.1180 . In addition, the proposed method showed strong potential to automate the PRODES project, which provides the official Cerrado yearly deforestation maps based on visual interpretation.

Keywords: deforestation; Cerrado; Brazilian savanna; time series; LSTM; U-Net; Landsat; Sentinel



Citation: Matosak, B.M.; Fonseca, L.M.G.; Taquary, E.C.; Maretto, R.V.; Bendini, H.d.N.; Adami, M. Mapping Deforestation in Cerrado Based on Hybrid Deep Learning Architecture and Medium Spatial Resolution Satellite Time Series. *Remote Sens.* **2022**, *14*, 209. <https://doi.org/10.3390/rs14010209>

Academic Editor: Sandra Eckert

Received: 10 August 2021

Accepted: 21 September 2021

Published: 3 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cerrado has huge importance for species conservation and ecosystem services, such as water and carbon, but it is highly threatened by deforestation. Thus, highlighting its importance, we divided the introduction into sections as follows: in Section 1.1, we give a brief introduction to the Cerrado biome in Brazil. In Section 1.2, we describe some Cerrado vegetation monitoring projects. In Section 1.3, we present techniques for mapping vegetation cover changes using artificial intelligence methods. Finally, in Section 1.4 we present our research scope.

1.1. Cerrado Biome in Brazil

Cerrado is the second largest biome in the Brazilian territory, with an area of approximately 2 million km². This biome composes 23% of the national territory, and covers areas of 11 states and the Federal District. With more than 4800 endemic species of plants and vertebrate animals, Cerrado is considered one of the global hotspots for biodiversity conservation, as it is under severe human-induced threats [1,2]. Aside from its biodiversity richness, this biome is crucial for the country's water supply due to the presence of headwaters and springs that originate rivers of important Brazilian watersheds [3,4]. The

Cerrado biome contains highly complex gradients of natural vegetation with important differences in herbaceous, woody, and forest layers, and its phenology is mostly related to the amount of water available in the soil [5,6]. This structural diversity of vegetation types encompasses a broad spectrum of biomass quantities, which is a large carbon stock and a source of CO₂ emissions during the conversion process of natural vegetation to agriculture and pasture areas [7].

Despite its large species diversity and ecological importance, Cerrado has presented high degradation rates since 1960 [1,8]. The conversion of natural vegetation to anthropic areas occurs at high rates, and more than 50% of its natural vegetation has already been converted, mainly into agriculture and pasture [9–11]. Studies have shown that it is possible to increase agriculture production in Cerrado through agricultural intensification and sustainable practices, among other actions to protect the remaining natural vegetation [12,13]. These practices should be encouraged by public environment conservation policies. However, to correctly direct those policies, it is necessary to accurately monitor the Cerrado native vegetation conversion to understand its land occupation dynamics [13].

1.2. Monitoring Cerrado Vegetation

Attempts to monitor deforestation and forest degradation in the Cerrado are relatively recent, unlike those for the Amazon, which began in 1988 [14,15]. Some initiatives started to monitor vegetation in the 2000s, with the Conservation and Sustainable Use of Brazilian Biological Diversity Project (PROBIO) mapping Cerrado's vegetation cover [16], along with deforestation alerts created by the Integrated System of Deforestation Alerts (SIAD) [17]. Cerrado deforestation maps were produced for the years 2010–2015 by the National Institute for Space Research (INPE), which were the basis for submitting a request for payments by avoided emissions. The production of these maps received financial support from the Ministry of Science, Technology and Innovation (MCTI), Ministry of the Environment (MMA) and the World Bank in addition to the German institutions Credit Institute for Reconstruction (KfW) and German Corporation for International Cooperation (GIZ) [18]. In 2016, Brazil submitted the request to the United Nations Framework Convention on Climate Change (UNFCCC) as a first action for the biome in the implementation of the REDD+ policies [19]. Based on this submission, MCTI had approved the project "Development of Forest Fire Prevention Systems Vegetation Cover Monitoring in the Brazilian Cerrado" by the World Bank [20,21]. This project, called FIP Monitoring, is part of the Brazilian Investment Plan (BIP) under the Forest Investment Program (FIP) [22]. With FIP financial support, the National Institute for Space Research (INPE; Portuguese acronyms) started to produce yearly deforestation maps for Cerrado through the Satellite Deforestation Monitoring Project (PRODES) and Near Real-time Deforestation Detection (DETER) in 2016 [15,23].

PRODES Cerrado monitors deforestation in the entire biome and provides yearly deforestation maps and deforestation increment rate. PRODES uses Landsat-like imagery and has an overall accuracy around 94% [24]. However, the PRODES deforestation detection procedure is performed by visual interpretation, which involves various remote sensing specialists in a laborious, high financial cost, and time-consuming task. These drawbacks have instigated researchers to develop semi-automatic methods to map deforestation in the Amazon and Cerrado, the two biggest biomes in Brazil [25,26]. Deforestation data provided by PRODES and DETER constitute powerful information to support the development of automatic methods to map deforested areas. Some initiatives have been launched to automate this process.

1.3. Mapping Vegetation Cover Changes Using Artificial Intelligence Techniques

Due to the heterogeneous and seasonal natural vegetation in Cerrado, it is a challenge to automate Cerrado Land Use and Land Cover (LULC) mapping and its changes [5,27–30]. Recent advances in remote sensing and artificial intelligence techniques, as well as the increasing amount of freely-available Earth Observation satellite imagery have allowed

the development of LULC change detection techniques in complex environments such as Cerrado [25,30–33].

Remote sensing time series derived from a sequence of images can provide important information to investigate the dynamics of the environment over time. Specifically for vegetation, phenology stages can be measured by the spectral response, such as budbreak, leaf out bloom, and leaf senescence of forest. These phenology stages can be associated with patterns extracted from time series, which can be used for LULC classification and change detection [34–36]. Methods such as Breaks For Additive Season and Trend (BFAST) Lite [37] and Jumps Upon Spectrum and Trend (JUST) [38] have provided interesting results for detecting changes using remote sensing time series. JUST can simultaneously search for trends and statistically significant spectral components of each time series segment to identify the potential jumps by considering appropriate weights associated with the time series. It is a robust change detection method which does not require any interpolations and/or gap fillings. According to [38], JUST is more resistant to a poorer signal-to-noise ratio than the change detection using the BFAST.

Although studies have been using Deep Learning (DL) algorithms for a variety of remote sensing tasks for the past few years, they are still relatively unexplored for deforestation mapping [39]. Recently, DL methods have shown promising results for deforestation mapping, with high accuracy, robustness to various sources of noise, and the ability for large-scale mapping [25,33,39]. To assess state-of-art pattern recognition methods, Adarme et al. [40] evaluated three DL techniques for automatic deforestation detection in the Brazilian Amazon and Cerrado biomes. The authors used two Landsat 8 images acquired at different dates. The strategies based on DL achieved the best performance in comparison with other methods and achieved an overall accuracy up to 95% for Cerrado. Similarly, Maretto et al. [33] developed a method based on DL U-Net architecture, incorporating both spatial and temporal contexts for detecting deforestation in the Brazilian Amazon regions using Landsat-8/OLI images. Using a real-world dataset, their method outperformed a traditional U-Net architecture, achieving approximately 94% overall accuracy when applied to the entire Pará State in Brazil. Furthermore, the authors adapted the method to apply it in a region comprising approximately 130,000 km² in the east of the Cerrado biome, in which an overall accuracy of approximately 92% was achieved. Another change detection approach using U-Net was proposed by [39]. In this case, the U-Net input is a stack created with two Landsat-8/OLI images, with one year of interval between their acquisition dates. Similar to [33], the changes are detected by comparing images acquired before and after the change events. Their method outperformed two machine learning algorithms outside of the DL scope in a direct comparison. However, the presence of clouds in the images can hinder the change detection process performed by this method.

When using DL together with image time series, classification methods can take advantage of temporal and spatial information to better discriminate classes with similar spectral information [35,41–44]. Within this context, Taquary et al. [45] proposed a method to detect Cerrado deforestation by combining two different DL architectures: the Long Short-Term Memory (LSTM) [46,47] to analyze temporal patterns, and the U-Net [48] to analyze spatial patterns. They used a time series composed of monthly composites of PlanetScope images with spatial resolution of 3 m. Despite the great results obtained in [45], the use of high resolution images imposes limitations due to the high cost of images as well as the high processing time due to the huge amount of data to cover the entire Cerrado biome. In addition, medium spatial resolution satellites such as Landsat and Sentinel provide cost-free images and present a good temporal resolution of 16 days and 5 days, respectively, [49,50]. Current LULC mapping projects have used these data [11,51] and integrated them in data cubes to generate dense image time series [52].

1.4. Research Scope

Considering all aspects related to change detection in complex environments, the main objective of this study is to develop a method to detect deforestation in the Cerrado biome

based on the combination of LSTM and U-Net techniques and time series derived from Landsat-8 and Sentinel-2 imagery. The hybrid classification based on LSTM and U-NET can produce deforestation maps faster than end-to-end DL architectures, which analyze time and spatial patterns at the same time, such as the ConvLSTM method [53,54]. Moreover, the time and spatial analysis performed in two steps allows the analysis based on more contextual information extracted from larger neighborhood areas, which can provide better classification results. Although our strategy is similar to the one proposed by [45], we carried out various improvements to adapt the methodology to medium spatial resolution imagery, which can be combined with such auxiliary data as slope information to improve the results. In addition, PRODES data, that is, the official data of the Brazilian government to implement the REDD+ policies, was used as reference to generate training samples.

This rest of the paper is organized as follows: Section 2 describes the study areas, methodology, and validation tests for detecting changes in Cerrado's native vegetation. Section 3 presents the results obtained for the study areas. Finally, Sections 4 and 5 present discussions and conclusions, respectively.

2. Materials and Methods

2.1. Study Areas

To test and validate the method, we defined two test sites, one in the state of Bahia and the other in the state of Mato Grosso, with approximately 35,200 km² and 27,500 km², respectively, (Figure 1). Each study area has two sub-areas: 'main' and 'auxiliary'. The first one is the area to be mapped and the second one is the region in which the training samples will be selected in some tests (Section 2.3.1).

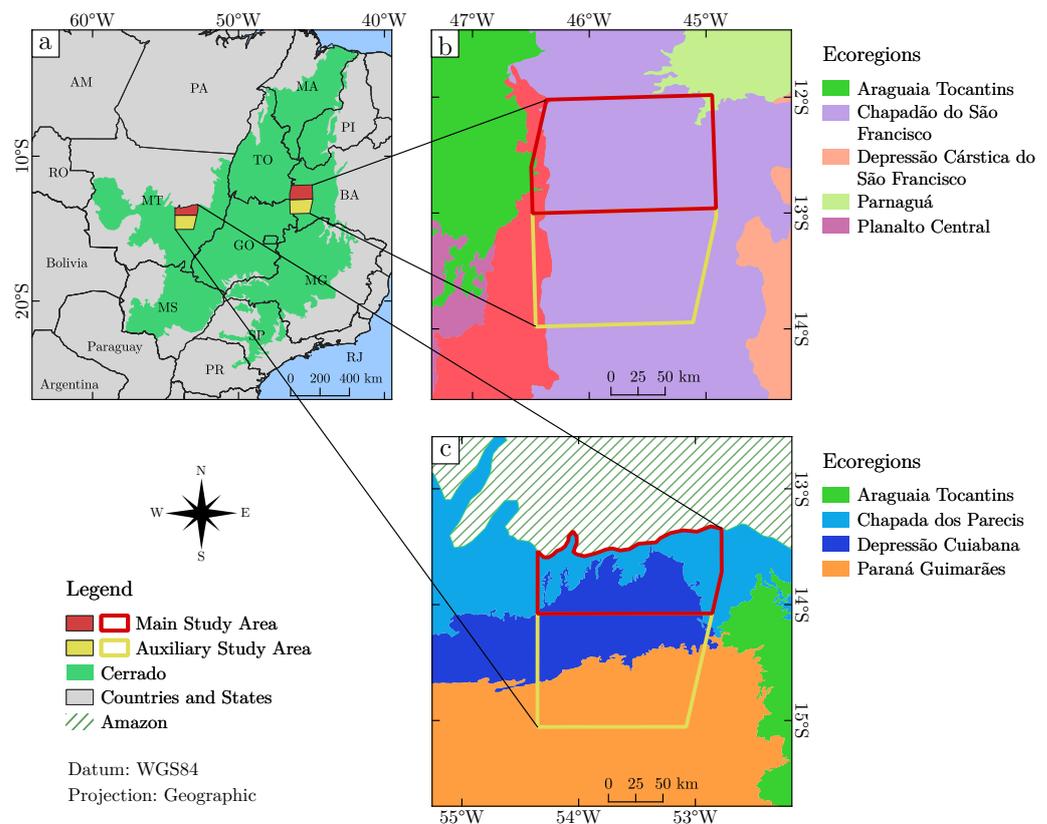


Figure 1. (a) Study area localization. (b) Bahia and (c) Mato Grosso. Ecoregions source: [13].

These regions were chosen because they contain the main deforestation patterns present in Cerrado. In the Mato Grosso study area, the agriculture and pasture are more consolidated, and deforestation has already affected areas that are suitable for mechanized agriculture. In this scenario, current deforestation tends to occur in smaller and more

protected areas with accidented topography. In the Bahia study area, the main land use is agriculture that occupies large fields due to terrain aptitude for mechanization. Agriculture expansion is responsible for current high deforestation rates in this region, which advances quickly in large and geometrical fields.

Sano et al. [13] proposed a division of Cerrado into different ecoregions that reflect the environmental heterogeneity within the biome. The Mato Grosso study area belongs to 3 ecoregions: Paraná Guimarães, Depressão Cuiabana, and Chapada dos Parecis (Figure 1c), granting it a higher level of pattern complexity due to its high heterogeneity. In addition, this study area is situated in a transition area between the Cerrado and Amazon biomes. On the other hand, the Bahia study area belongs mainly to the Chapadão do São Francisco ecoregion (Figure 1b) and is located far from the biome borders.

The study areas also present different deforestation patterns. In Bahia, deforestation polygons occur in large geometrical fields, while in Mato Grosso they occur in a great number of amorphous polygons. For 2019, the number of PRODES polygons in Mato Grosso is higher, whilst the total deforestation area and the mean area per polygon are higher for Bahia, as can be observed in Table 1.

Table 1. Deforestation statistics for 2019 in the study areas (main and auxiliary). Source: adapted from [11].

Deforestation (2019)	Mato Grosso	Bahia
Total Area (ha)	13,326.516	20,723.315
Mean Polygon Area (standard deviation; ha)	18.93 (± 55.38)	61.68 (± 235.60)
Polygon Count (unit)	704	336
Percentage of Study Area (%)	0.484	0.589

2.2. Input Data

Cerrado has potential for arable land use, and technological advances in soil and crop management have made it a suitable region for mechanized agriculture [9]. Consequently, deforestation has occurred mainly in large areas in Cerrado. Thus, we included in the processing the terrain slope information derived from the Shuttle Radar Topography Mission (SRTM) [55], with a spatial resolution of 30 m. In the processing step, it was upsampled to match the pixel size of the image time series.

We used the PRODES Cerrado data [11] as the training sample reference, which were acquired in two different datasets: deforestation data that occurred prior to 2000 and after 2000, separated by year. Afterwards, these vectors were merged and used to generate datasets for the years 2018 and 2019 for each study area. PRODES produces yearly deforestation maps from images acquired in the dry season to avoid the interference of clouds. This approach defines the PRODES year, i.e., the deforestation reported in PRODES 2018 corresponds to deforestation detected from August 2017 to July 2018. This type of time interval defined by PRODES was taken into account in the generation of the image time series.

Two types of dense image time series were generated, one from Landsat-8/OLI and the other from Sentinel-2/MSI. To avoid spectral information differences between both satellites, we considered only similar bands to apply the method, as shown in Table 2, and also the Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) [38]. The satellite images were obtained from data cubes generated by the Brazil Data Cube project (BDC) (<http://brazildatacube.org/>, accessed on 10 June 2021) [56]. Data cubes of Landsat-8/OLI and Sentinel-2/MSI images have pixel sizes of 30 m and 10 m, respectively, even for Bands 8a, 11, and 12 of Sentinel, which were upsampled to 10 m. Together with Landsat and Sentinel data cubes, BDC provides Cloud and cloud shadow masks derived from FMask 4.2 [57]. All images available from August 2017 to August 2019 were acquired for both study areas. They were used to create temporal stacks comprising approximately one year, from August 2017 to August 2018 (Year 2018) and from August 2018 to August 2019 (Year 2019, similarly to the PRODES project). The Landsat-8/OLI

and Sentinel-2/MSI time series were created with a temporal resolution of 16 and 5 days, respectively. Missing values caused by clouds, cloud shadows, and unavailable images were estimated using a CubicSpline algorithm [58,59].

Table 2. Landsat-8/OLI and Sentinel-2/MSI time series data used as input t LSTM. Source: [60].

Landsat-8/OLI		Sentinel-2/MSI		Description
Name	Wavelength (μm)	Name	Wavelength (μm)	
Band 2	0.452–0.512	Band 2	0.458–0.523	Blue, surface reflectance.
Band 3	0.533–0.590	Band 3	0.543–0.578	Green, surface reflectance.
Band 4	0.636–0.673	Band 4	0.650–0.680	Red, surface reflectance.
Band 5	0.851–0.879	Band 8a	0.855–0.875	Near Infrared (NIR), surface reflectance.
Band 6	1.566–1.651	Band 11	1.565–1.655	Short Wave Infrared 1 (SWIR 1), surface reflectance.
Band 7	2.107–2.294	Band 12	2.100–2.280	Short Wave Infrared 2 (SWIR 2), surface reflectance.
NDVI	–	NDVI	–	Normalized Difference Vegetation Index.
EVI	–	EVI	–	Enhanced Vegetation Index.

2.3. Deforestation Detection

The flowchart in Figure 2 summarizes the methodology used to detect deforestation in Cerrado. Combinations of image time-series types, approaches to select training samples, and study areas were evaluated.

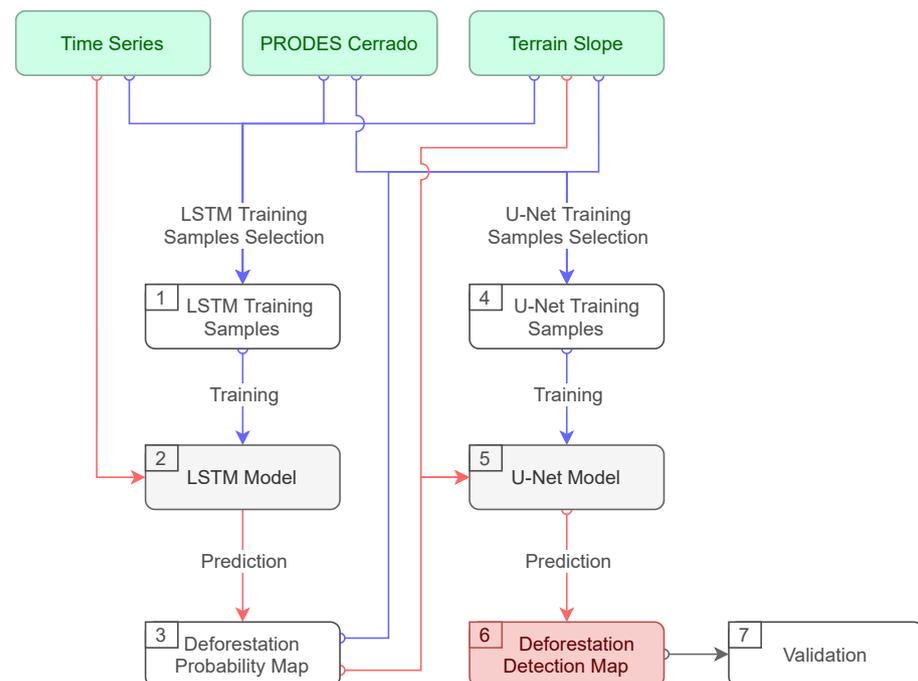


Figure 2. Flowchart to describe the deforestation detection methodology. Rectangles: the input data, DL models, and final results are colored in green, gray, and red, respectively. The numbers indicate their order in the flowchart. Connections: operations regarding the models training, operations related to using the trained models to predict over data, and operations of validation are colored in blue, red, and gray, respectively.

The data sources were used to select training samples for the LSTM algorithm. After the training phase, the LSTM was applied to the satellite image time series to create the deforestation probability map, which takes into account information in the time axis of each image pixel. The deforestation probability map was then employed with PRODES and Slope data to generate training samples for the U-Net algorithm. After training, the U-Net

produces the deforestation map, taking into account the spatial information provided by the probability map and terrain slope.

Afterwards, 12 deforestation maps were created considering all possible combinations of 2 satellite image time series (Landsat-8/OLI or Sentinel-2/MSI), 3 training samples scenarios (Approaches 1, 2, and 3), and 2 study areas (Mato Grosso and Bahia). These maps were then validated using a random sampling approach and reference data selected by visual interpretation. Three LULC classes were considered in our study (Table 3) Deforestation, Natural Vegetation, and Past Deforestation. Since the Cerrado biome is composed of Forest, Savanna, and Grassland formations, the term Deforestation will be associated to the suppression of any type of Cerrado vegetation in this research. The dynamics of the Cerrado natural vegetation are usually associated with fires and play a fundamental role in the ecological functioning of this biome. In this case, burned areas are not considered as Deforestation in this study, but as a type of degradation. However, the natural fire regime has been altered by anthropic land-use practices. Fires used by farmers to induce the regrowth of pastures during the dry season tend to get out of control, spreading over large areas and affecting areas of environmental protection, indigenous lands, and remnants of natural vegetation [61,62]. Other methods have been used to identify fire scars [63,64].

Table 3. Mapping classes in deforestation detection.

Class	Description
Deforestation	Total natural vegetation removal (change) caused by human activity in 2019.
Natural Vegetation	Natural vegetation (no change) during 2019.
Past Deforestation	Deforestation detected by PRODES before 2019 (masked).

Figure 3 shows a noticeable difference between natural vegetation and deforestation pixel time series. The natural vegetation NDVI starts to rise in November 2018 during the wet season, it reaches a plateau in January 2019, and starts to slowly decrease in June 2019 during the dry season. On the other hand, the deforestation NDVI profile follows a similar pattern as the natural vegetation in the beginning, but thereafter it presents an abrupt fall in April 2019, which is related to the natural vegetation suppression. This pattern observed in the time series can be used to describe changes over time.

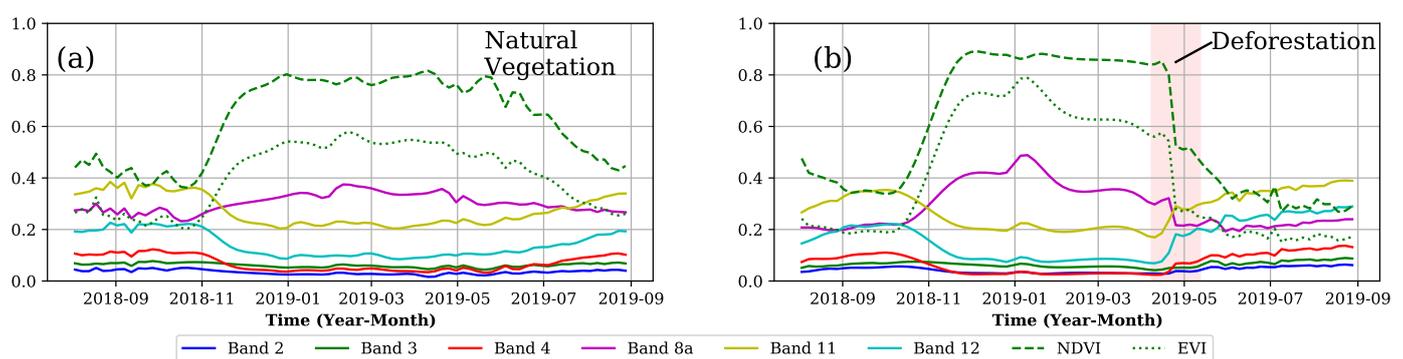


Figure 3. Examples of Sentinel LSTM training samples time series for (a) Natural Vegetation and (b) Deforestation.

2.3.1. Approaches for Training Samples Selection

The training samples selection is as important as the detection algorithm architecture [65,66]. The most difficult and time-consuming task in the classification process is the reference data generation. For this task, we evaluated 3 different approaches to select training samples, considering data variations in space and time. These strategies were implemented to better understand the limitations that may exist in each case, as well as

to explore the transferability of the models trained with samples obtained in different locations and time periods. In Approach 1, training samples were selected in the same main study area in which the classification is applied. In this case, areas used to select training samples were excluded during the process of validation. In Approach 2, training samples were selected in the ‘auxiliary study area’, an adjacent region to the main study area in which the algorithm is applied. The ‘auxiliary study area’ is highlighted in yellow, as shown in Figure 1. In Approach 3, training samples were selected in the main study area for images acquired in 2018, while the classification is carried out in the main study area for images acquired in 2019. Figure 4 illustrates the three samples’ training strategies described above.

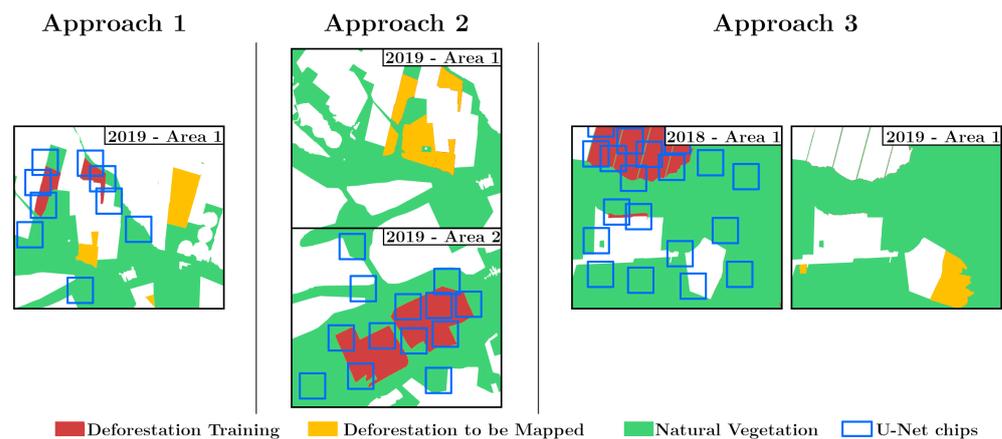


Figure 4. Strategies to select training samples. White colored areas represent past deforestation.

2.3.2. Long Short-Term Memory Training and Prediction

The first step of the hybrid DL classification consists of applying an LSTM to evaluate the time series in the time axis without considering the spatial context. This architecture was implemented in the Python programming language using the Tensorflow [67] DL library. It is built in 3 layers: (1) one LSTM layer with 256 hidden units, *tanh* activation function, and *sigmoid* recurrent activation function; (2) one batch normalization; and (3) one fully-connected output layer (dense) with *softmax* activation function. This model was trained with batches of 256 samples, *Adam* optimizer, and the loss function *Categorical Cross-Entropy*. The learning rate and number of epochs were empirically and individually optimized for each LSTM model, in order to achieve the highest accuracy without overfitting or underfitting. The learning rate ranged between 5×10^{-5} and 1×10^{-6} , and the number of epochs was between 1000 and 5000. In the training process, we used samples with shape $[e, b]$, where e is the amount of time entries in the time series and b is the number of input data (bands and vegetation indices). Each LSTM training sample is composed of the time series of a pixel, with the bands and vegetation indices described in Table 2. The same amount of deforestation and natural vegetation samples were selected from PRODES data to avoid class imbalance problems [66]. Deforestation samples were randomly selected inside PRODES deforestation polygons, whose amount in each polygon is limited to avoid overrepresenting the pattern of large polygons to the detriment of small ones. A similar principle was performed to select natural vegetation samples, but in this case the samples were stratified according to the topography slope since deforestation and vegetation physiognomies in Cerrado are correlated with topography [9,68].

The LSTM returns values between 0 and 1 that indicate the probability of a pixel time series belonging to the deforestation class. After training, the model is used in a prediction procedure for the remaining study area. The LSTM result is a deforestation probability map, which is used in the next processing phase carried out by U-Net algorithm.

2.3.3. U-Net Training and Prediction

The LSTM output maps contain noise and some regions of unclear deforestation occurrence. We used a U-Net to evaluate spatial patterns in the LSTM deforestation probability map and topography slope. The U-Net algorithm was implemented in the Python programming language with DeepGeo [69], a library based on TensorFlow [67]. Hyperparameters such as epochs, learning rate, and decay rate were optimized during each training, but common parameters were: a sample size of 284×284 pixels; loss function *Average Soft Dice*; learning rate decay activated; L2 regression rate of 0.0005; and 6 data augmentation operations per sample (rotation in 90° , 180° , and 270° , and flip horizontally, vertically, and transpose).

To train the U-Net model, samples were composed of LSTM deforestation probability, slope, and PRODES reference, as illustrated in Figure 5. Each sample is called ‘chip’ and all of them have the same dimensions. Since Landsat and Sentinel chips have the same size of 284×284 pixels, a Landsat chip covers a larger area than a Sentinel chip due to its larger pixel size. Consequently, a Landsat-8 chip contains more context information but less detailed information than the Sentinel-2 chip due to the better spatial resolution of the Sentinel-2 satellite.

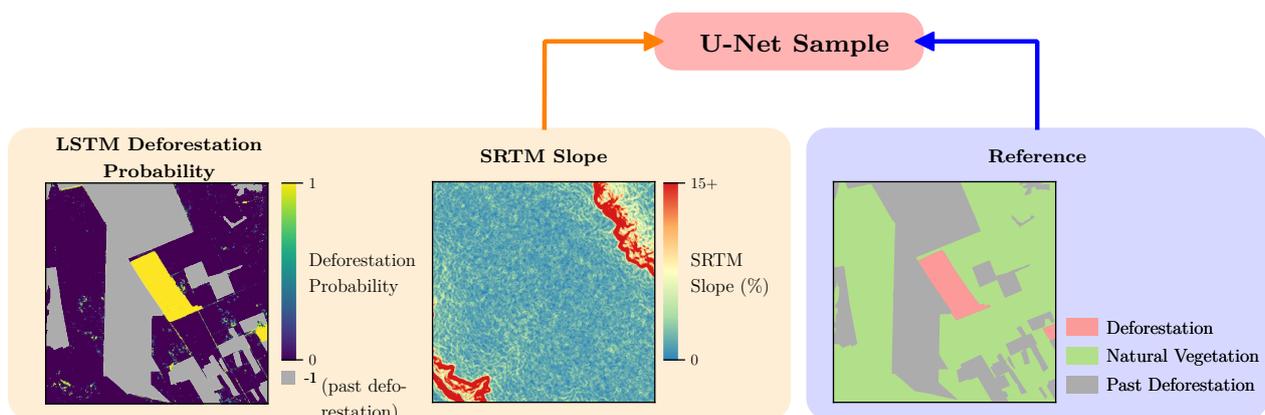


Figure 5. U-Net sample (chip): LSTM map, SRTM slope, reference deforestation. Samples have the same dimensions of 284×284 pixels.

2.4. Validation

In the validation process, it is recommended to use reference data of a better spatial resolution [70]. Since PRODES deforestation maps are based on Landsat-like images, we used Sentinel-2 time series to validate all maps produced in our study. The validation procedure was based on a stratified random sampling approach. The number of validation points (n) was defined by the following equation:

$$n = \frac{z_{\alpha/2}^2 \cdot \sigma^2 \cdot N}{e^2(N-1) + z_{\alpha/2}^2 \cdot \sigma^2} \quad (1)$$

We considered a variance of 50% ($\sigma^2 = 0.5$), a standard error of 3% ($e = 0.03$), a confidence interval of 95% ($z_{\alpha/2} = 1.96$), and the population size (N) [24,71]. This procedure resulted in 1067 validation points per map. As the deforestation area in our study regions is significantly smaller than the natural vegetation one, we defined 100 points stratified in deforestation as recommended in [70] to avoid under-representation of the deforestation class. On the other hand, 967 points were stratified in natural vegetation. The probability of each stratified point in relation to its category is shown in Table 4. Reference data for every validation point was independently created by visual interpretation over Sentinel-2 time series in the study areas. A total of 12,804 validation points were created to represent *deforestation* and *natural vegetation* classes.

Table 4. Sample probabilities for the validation process.

Study Area	Time Series	Samples	Class	Population (Pixels)	Samples Probability	Samples Weight	Sample Size	
							Map	Strata
Mato Grosso	Landsat	Approach 1	Deforestation	40,663	0.002459	407	1067	100
			Natural Vegetation	6,749,089	0.000143	6979		967
		Approach 2	Deforestation	41,997	0.002381	420	1067	100
	Natural Vegetation		6,473,917	0.000149	6695	967		
	Approach 3	Deforestation	53,868	0.001856	539	1067	100	
		Natural Vegetation	6,462,098	0.000150	6683		967	
	Sentinel	Approach 1	Deforestation	359,413	0.000278	3594	1067	100
			Natural Vegetation	61,159,707	0.000016	63,247		967
		Approach 2	Deforestation	147,986	0.000676	1480	1067	100
Natural Vegetation	61,369,681		0.000016	63,464	967			
Approach 3	Deforestation	264,001	0.000379	2640	1067	100		
	Natural Vegetation	61,255,249	0.000016	63,346		967		
Bahia	Landsat	Approach 1	Deforestation	64,635	0.001547	646	1067	100
			Natural Vegetation	9,125,501	0.000106	9437		967
		Approach 2	Deforestation	72,357	0.001382	724	1067	100
	Natural Vegetation		9,114,557	0.000106	9426	967		
	Approach 3	Deforestation	92,527	0.001081	925	1067	100	
		Natural Vegetation	8,853,307	0.000109	9155		967	
Sentinel	Approach 1	Deforestation	725,517	0.000138	7255	1067	100	
		Natural Vegetation	84,341,212	0.000011	87,219		967	
	Approach 2	Deforestation	636,114	0.000157	6361	1067	100	
Natural Vegetation		84,428,617	0.000011	87,310	967			
Approach 3	Deforestation	543,672	0.000184	5437	1067	100		
	Natural Vegetation	84,521,308	0.000011	87,406		967		

In the validation process, the confusion matrix for each map was used to calculate the Overall Accuracy and F1-Score. Their confidence interval was obtained through $\pm Z_{\alpha/2}(SE)$, where $\alpha = 95\%$ ($Z_{\alpha/2} = 1.96$). For Overall Accuracy, the SE value was calculated as demonstrated in [70]. For the F1-Score measure, its SE was obtained through propagation of Precision and Recall standard errors, since F1-Score is a function of them. The standard errors for the Precision and Recall were obtained according to [70], in which they are referred as Producer's and User's accuracy.

3. Results

Twelve deforestation maps were produced through the combination of two study areas (Mato Grosso and Bahia), two satellite image time series (using Landsat-8/OLI and Sentinel-2/MSI sensors), and three training sample selection approaches. Figure 6 shows the LSTM deforestation probability map and the U-Net final result map for the Mato Grosso study area using Landsat time series and sample selection based on Approach 3. Various locations in the probability map have probability values around 0.5 with some spurious points. The U-Net algorithm filtered this noise and also preserved the polygons in which deforestation occurred despite the presence of noise.

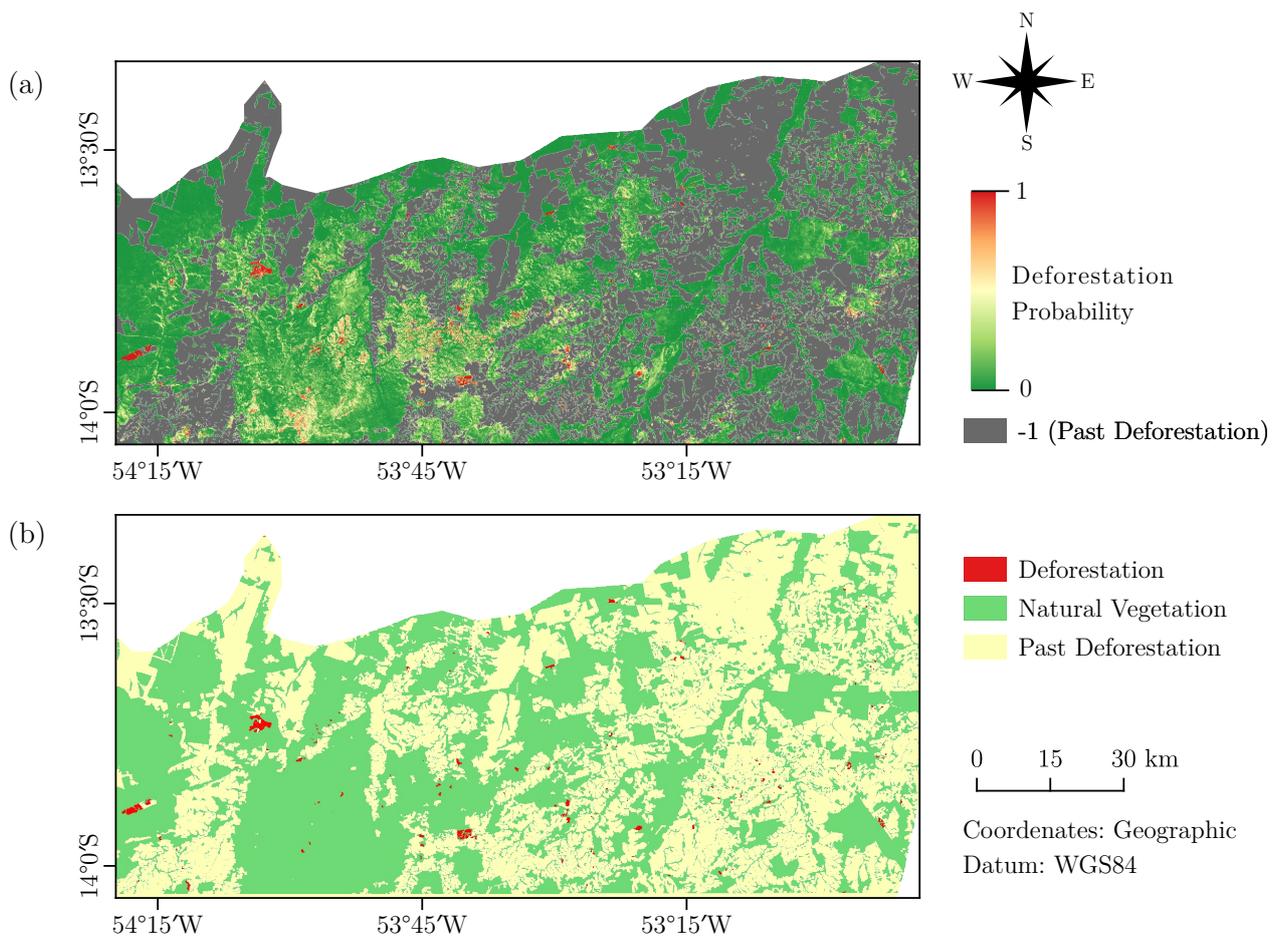


Figure 6. Results for Mato Grosso study area, Landsat time series and training samples based on Approach 3. (a) LSTM deforestation probability map; (b) U-Net deforestation map.

Figure 7 shows the LSTM deforestation probability map and the U-Net deforestation map for the Bahia study area using Landsat time series and training samples based on Approach 3. In this case, the deforestation probability maps were more precise and presented less noise compared to the Mato Grosso study area. However, the presence of probability values around 0.5 is still noticeable.

Figure 8 shows inset maps for the Mato Grosso study area with focus on the largest deforestation polygon detected in this area. We can observe that the most accurate deforestation corresponds to the Approach 1 with Sentinel-2/MSI data. This may be due to the higher spatial and temporal resolutions of Sentinel-2/MSI, as well as to the higher correlation degree between training samples and the mapped region. In this case, the training samples were selected in the 'main' study area and in the same year of 2019 as the resultant deforestation map. Although this result is very similar to the PRODES deforestation, we can observe some noise not present in the PRODES and Sentinel-2 images. On the other hand, the deforestation maps obtained through Approach 2 have inferior quality in relation to the others. This could be caused by the the 'auxiliary' study area used to select the training samples that contains different ecoregions (Figure 1). Consequently, the input data and the training area are not well correlated, which demonstrate that differences in vegetation and soil types can influence the deforestation detection process, impacting the transferability of the classification model.

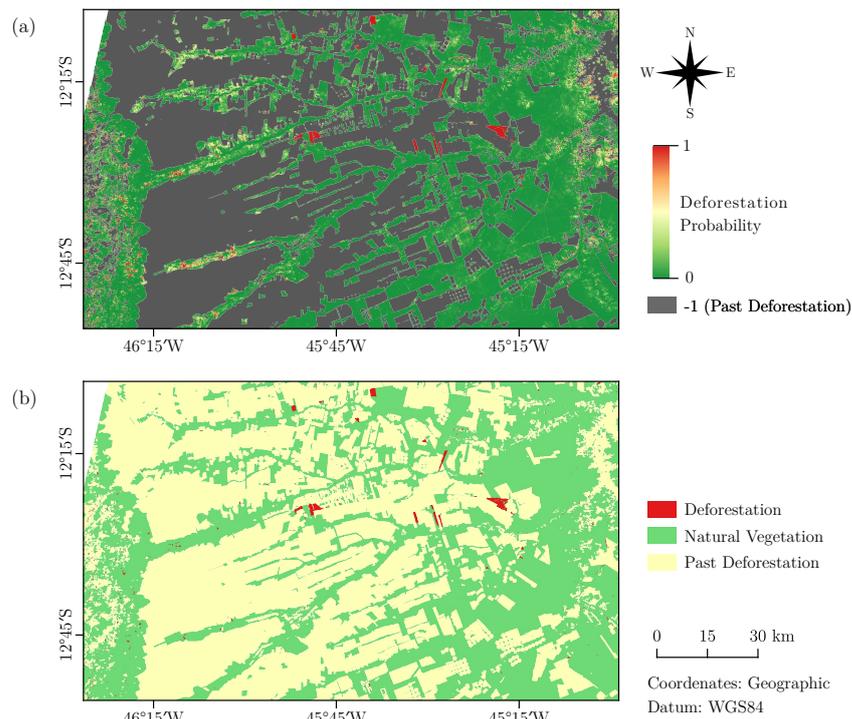


Figure 7. Results for Bahia study area using Landsat time series and training samples based on Approach 3. (a) LSTM deforestation probability map; (b) U-Net deforestation map.

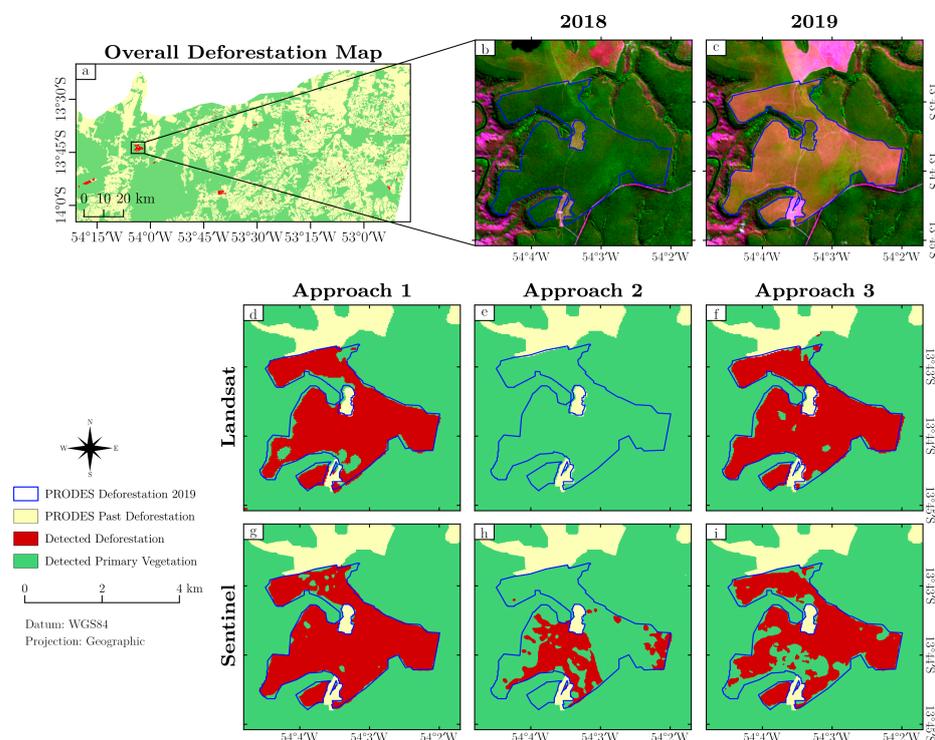


Figure 8. Results obtained for the largest deforestation polygon in the Mato Grosso study area. (a) PRODES Deforestation in 2019 in the study area in Mato Grosso. (b,c) RGB (Band 11, Band 8, Band 4) Sentinel-2/MSI images in 4 July 2018 and 4 July 2019, respectively, overlaid on the largest PRODES deforestation polygon for Mato Grosso study area in 2019; (d,e,f) Deforestation maps using Landsat-8/OLI time series and Approaches 1, 2, and 3, respectively; (g,h,i) Deforestation maps using Sentinel-2/MSI time series and Approaches 1, 2, and 3, respectively.

Figure 9 shows inset maps of the results for Bahia, focusing on the largest deforestation polygon for this area. Visually, the most accurate deforestation mapping was provided with Sentinel-2/MSI data and Approaches 1 and 2 (Figure 9g,h). In these cases, confusion among classes was not present inside the deforestation polygons, like in Approach 3 (Figure 9f,i). Furthermore, the proposed method correctly detected thin natural vegetation corridors in the deforestation area, which was not successfully mapped in Approaches 1 and 2 using Landsat-8/OLI data (Figures 9d,e).

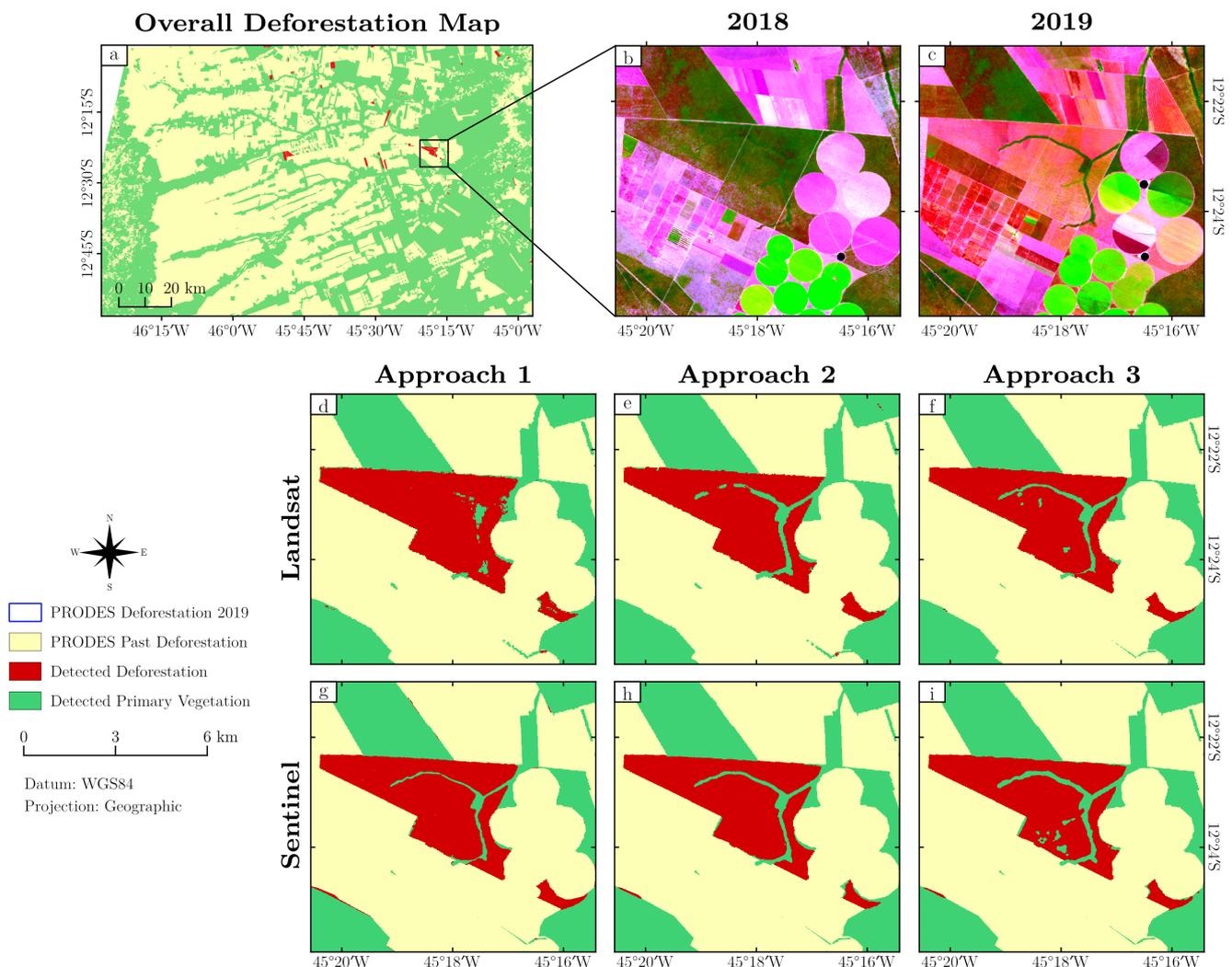


Figure 9. Results obtained for the largest deforestation polygon in the Bahia study area. (a) Deforestation map by PRODES for 2019 for the ‘main’ study area in Bahia. (b,c) RGB (Band 11, Band 8, Band 4) Sentinel-2/MSI images for 4 July 2018 and 4 July 2019, respectively, overlaid by the largest PRODES deforestation polygon for the Bahia study area in 2019; (d,e,f) Results using Landsat-8/OLI time series for Approaches 1, 2, and 3, respectively; (g,h,i) Results using Sentinel-2/MSI time series for Approaches 1, 2, and 3, respectively.

Table 5 shows accuracy metrics for all deforestation maps. The highest values belong to group *a* and the lowest belong to group *b*. The results suggest that the deforestation detection was more successful for the Bahia study area. For the Bahia study area, the use of Sentinel time series showed some superiority, with Approaches 1 and 3 performing similarly.

Table 5. Overall Accuracies for the maps and F1-Scores for their deforestation class. The exponent letters represent the groups in which the Overall Accuracies or F1-Scores are statistically the same. Group *a* comprises the highest values and group *b* the lowest.

Study Area	Time Series	Samples	Overall Accuracy	F1-Score
Mato Grosso	Landsat	Approach 1	99.58% ± 0.35 ^{ab}	0.7128 ± 0.1714 ^{ab}
		Approach 2	98.96% ± 0.57 ^b	0.4486 ± 0.1395 ^b
		Approach 3	99.60% ± 0.29 ^{ab}	0.6864 ± 0.1588 ^{ab}
	Sentinel	Approach 1	99.15% ± 0.53 ^{ab}	0.5428 ± 0.1570 ^b
		Approach 2	98.99% ± 0.60 ^b	0.5255 ± 0.1503 ^b
		Approach 3	99.09% ± 0.53 ^{ab}	0.4963 ± 0.1496 ^b
Bahia	Landsat	Approach 1	99.40% ± 0.45 ^{ab}	0.6954 ± 0.1608 ^{ab}
		Approach 2	99.81% ± 0.21 ^a	0.8739 ± 0.1183 ^a
		Approach 3	99.52% ± 0.35 ^{ab}	0.7091 ± 0.1531 ^{ab}
	Sentinel	Approach 1	99.81% ± 0.21 ^a	0.8795 ± 0.1180 ^a
		Approach 2	99.61% ± 0.35 ^{ab}	0.7768 ± 0.1559 ^{ab}
		Approach 3	99.77% ± 0.21 ^{ab}	0.8511 ± 0.1191 ^a

4. Discussion

4.1. Study Areas, Input Data and Training Samples

Considering the results for both study areas, we found that the deforestation was more accurately mapped in Bahia according to its higher F1-Scores (Table 1). This indicates that higher environmental complexity, present in the Mato Grosso study area, is a challenge in detecting Cerrado deforestation even for DL methods. Related literature also indicates that the discrimination between natural vegetation and other LULC becomes more difficult in Cerrado as the vegetation structure and canopy becomes more heterogeneous [29,72].

Another possible reason for better results in the Bahia study area is that the deforestation polygons, although fewer in number, have large areas. Conversely, the deforestation polygons in Mato Grosso have smaller areas and amorphous shapes and also occur in large numbers (Table 1). The smaller and more amorphous polygons caused the inclusion of mislabeled deforestation samples in the LSTM training phase in Mato Grosso. As PRODES has small uncertainties in polygons borders, the selection of time-series samples within small areas increases the chance of selecting samples at polygon borders. Nevertheless, mislabeled reference data spoil training samples and consequently degenerate the DL models performance [65,73]. In addition, as the complex deforestation polygons in Mato Grosso present higher uncertainties, the deforestation reference data were less accurate than the reference for Bahia [14].

Considering the results for the Bahia study area and Landsat data, the processing time took around 1 h to train the LSTM model, to create the deforestation probability map, to train the U-Net, and to generate the final deforestation map using a Tesla V100-SXM2-16GB GPU. Given that the study area is approximately 18,000 km², we estimated the processing time to map the entire Cerrado biome to be around 111 h. For this estimate, we considered the Cerrado biome divided into smaller regions with the same size as the Bahia study area, and also one LSTM and U-Net model for each region, independently of its complexity degree. In the case of partitioning the biome in larger areas, such as the ecoregions, the estimated time can be shorter and the deforestation mapping process can probably be less complex.

Two types of image time series were used in this work: Landsat-8/OLI and Sentinel-2/MSI. Figures 8 and 9 show the superiority of Sentinel-2/MSI over Landsat-8/MSI data when comparing the deforestation maps. According to [72], Landsat data present some limitations for Cerrado LULC mapping in regions that appear as a mosaic of grassland, savanna, and forest formations. Moreover, Lima et al. [74] stated that Sentinel-2 data presented better results than Landsat data for mapping selective logging in the Brazilian

Amazon region through a method based on time series. Bueno et al. [75] also indicate Sentinel data to achieve better LULC mapping results. The advantages of Sentinel imagery reported above can be explained by its better temporal and spatial resolution that can provide dense time series, which describe temporal patterns with more detail and also reduce the data scarcity due to clouds or cloud shadows [28,76]. These results stated here agree with our results, which pointed out better deforestation maps obtained using Sentinel data.

The analysis of three approaches to select training samples showed better results for Approaches 1 and 3. In the Mato Grosso study area, Approach 2 presented the worst results, with its F1-Scores belonging only to group *b* in Table 5. This result contrasts with Approach 2 for the Bahia study area, in which the F1-Scores belong to group *a*. In the case of Approach 2, the region in which training samples were selected and the region used to generate the deforestation map have different characteristics because they belong to different ecoregions in Mato Grosso (Figure 1). Therefore, their vegetation and deforestation patterns are different, which jeopardizes the training sample quality and consequently negatively impacts the classification model [65,66,73]. On the other hand, Table 1 shows that there are no significant differences between Approaches 1 and 3. In this case, training samples are selected from the prior year and in the same area as that in which the deforestation map was generated. We have to observe that spectral information differences due to images taken at different times can impact the classification results [77]. A strategy to minimize this influence consists of extracting information from more time periods to increase the training sample number in order to represent a wider variety of deforestation patterns. This matter is important because DL model accuracy depends on the training data quality and the pattern representation [40,78].

4.2. Comparison with PRODES

Parente et al. [24] performed the quality assessment of PRODES Cerrado data. They reported an Overall Accuracy of $93.17\% \pm 0.89$ for PRODES 2018. Although PRODES produces deforestation maps by visual interpretation, its methodology is robust and produces high accuracy maps for the entire Cerrado biome. Compared with PRODES, our method needs a larger amount of data and the data preparation and deforestation detection tasks require high performance computing. Nevertheless, it is semi-automatic and has the potential to be automated to the entire Cerrado biome as long as we can manage to optimize the training parameters for each ecoregion.

Considering the deforestation maps produced by the proposed methodology using Approach 1 and Sentinel-2/MSI image time series, Figure 10 shows its agreement with the PRODES Cerrado deforestation map. The agreement percentage for the deforestation and natural vegetation classes is about 99.81% and 99.67% for the Bahia and Mato Grosso study areas, respectively. The percentage values were obtained through dividing the agreement area by the total area in each study region. In both cases, the agreement is high, although PRODES data is produced using Landsat images (30 m). For the Mato Grosso region, the agreement is lower due to the region complexity since it is composed of different ecoregions.

We observe that PRODES and the deforestation maps produced by our method present a high agreement for both the Deforestation and Natural Vegetation classes. There was only one visible classification error in the Mato Grosso study area due to the subtle change pattern in its deforestation polygon. Figure 11 shows the LSTM deforestation probability for this polygon, which explains the classification error.

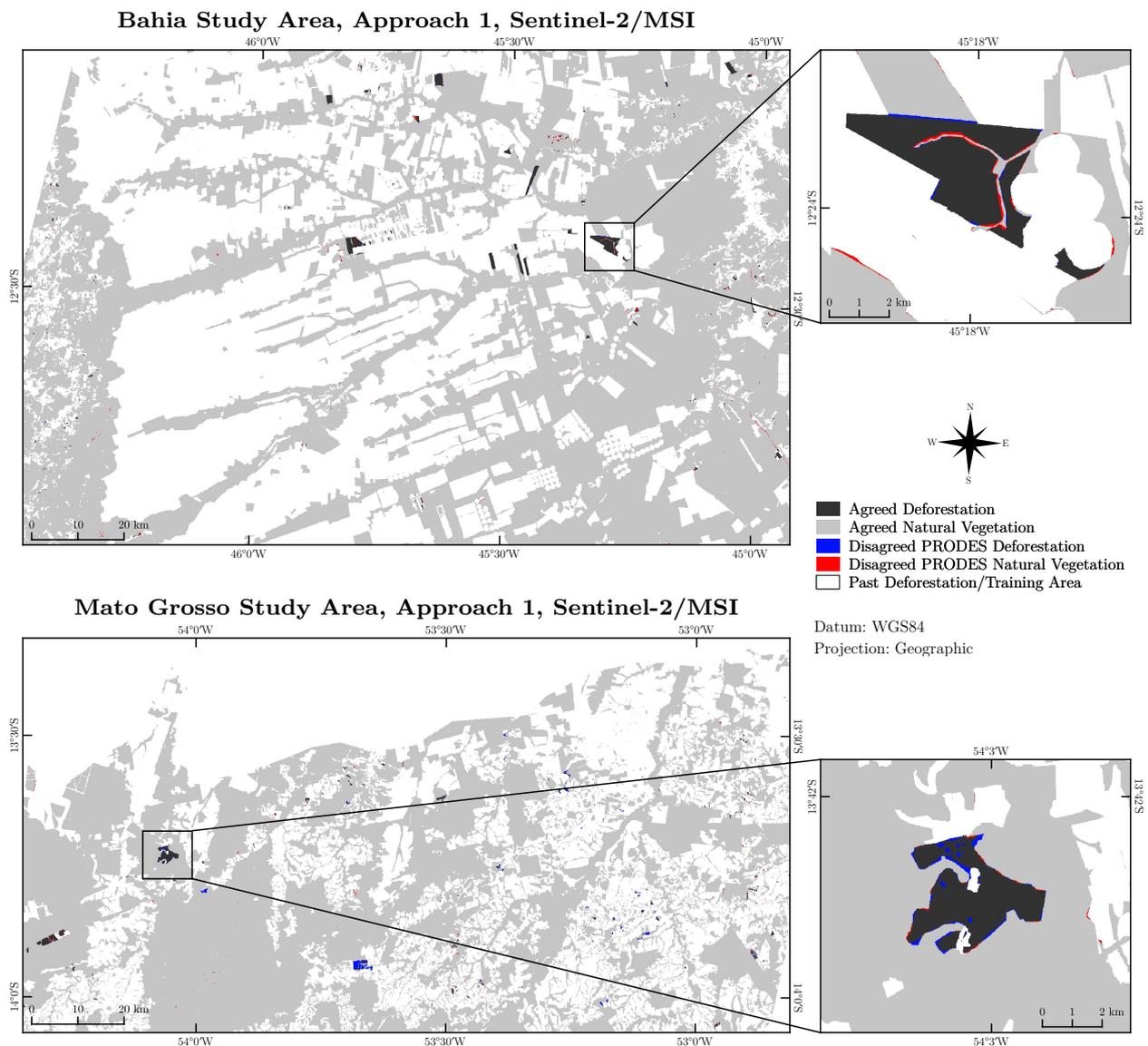


Figure 10. Agreement comparison between PRODES Cerrado 2019 and the deforestation maps using Sentinel-2/MSI and training samples based on Approach 1.

Figure 11a,b show the deforestation detected by PRODES. Figure 11c shows only parts of this polygon that obtained accurate high deforestation probability values. Comparing the deforestation probability values with natural vegetation in Figure 11a, we note that in locations where the natural vegetation was greener, the deforestation probability values were high, while the probability values were low in areas depicted in magenta. Through visual comparison between the Landsat images for the deforestation polygon and Google Earth high-resolution imagery for the year prior to the deforestation, we verified that magenta regions in the Landsat imagery in the deforestation polygon belong to the Grassland Formation while green regions belong to the Savanna Formation, as shown in Figure 11d,e. Therefore, the LSTM model produced less accurate deforestation polygons detected in Grassland Formation in the Mato Grosso region, hence disagreeing with PRODES. Other studies also reported difficulties for LULC mapping regarding Grassland Formation [27,28].

Natural wildfires are caused by lightning in Cerrado; however, human-started fires are common in the biome and are more intense, severe, large in size, difficult to control and costly for fighting [61]. PRODES defines deforestation as the clear cutting of natural vegetation, and it does not consider fire followed by natural vegetation regrowth as defor-

estation. Since we used PRODES data as the reference for our training samples, our results do not detect change associated with fire followed by natural vegetation regrowth.

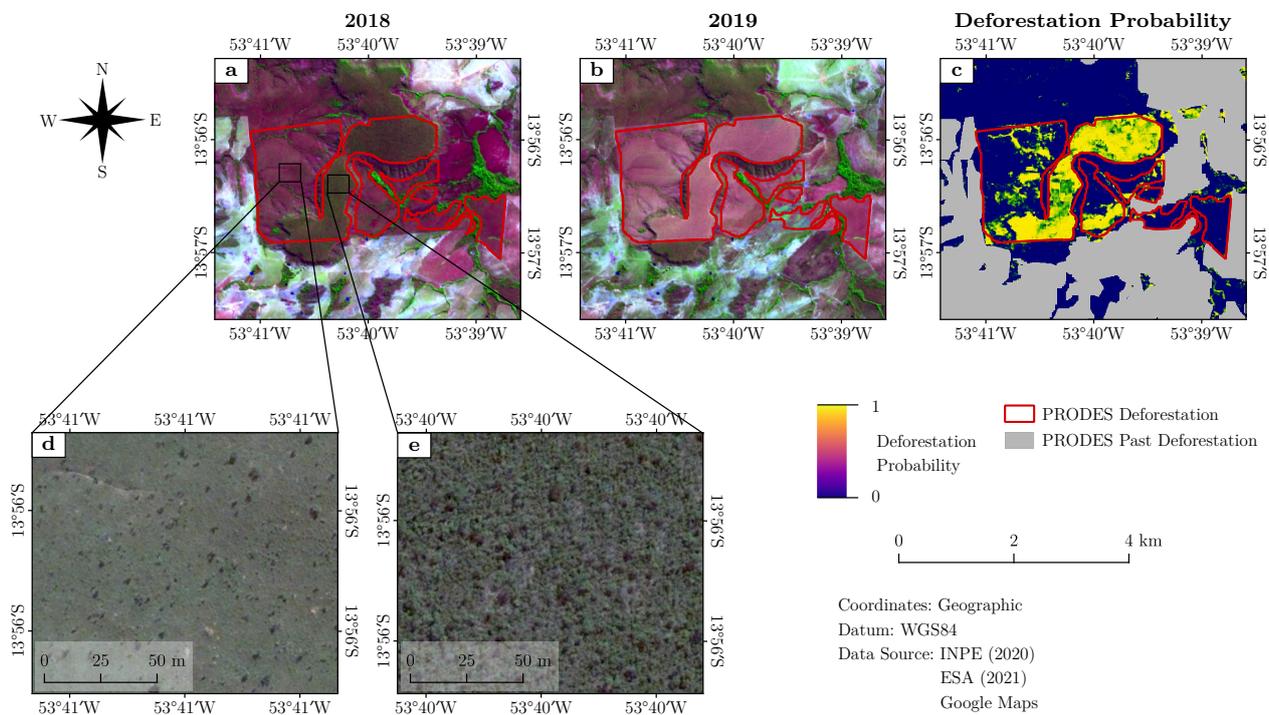


Figure 11. (a,b) RGB (Band 11, Band 8, Band 4) Sentinel-2/MSI images acquired on 3 July 2018 and 3 July 2019 (during the dry season), respectively, overlaid on PRODES deforestation polygon for the Mato Grosso study area (2019). (c) is the LSTM deforestation probability map overlaid on PRODES deforestation polygon and PRODES past deforestation. (d,e) are Google Earth images of natural vegetation during dry season prior to the deforestation in 2019, in which the first one is a Grassland Formation and the other is a Savanna Formation. LSTM deforestation probability maps are more accurate for Savanna Formations.

5. Conclusions

In this work, we proposed a methodology to detect deforestation in the Cerrado biome using Landsat and Sentinel time series through a combination of two DL architectures: LSTM and U-Net. The method uses PRODES deforestation polygons as a reference and image time series generated from Landsat and Sentinel-2 images to train the LSTM model. The probability map resulting from the LSTM is combined with PRODES and SRTM slope data to train the U-Net model, which is used to produce the final deforestation map.

The proposed method showed great potential to be applied to medium spatial resolution images such as Landsat-8 and Sentinel-2 to detect deforestation in Cerrado with a high overall accuracy of $99.81\% \pm 0.21$. The combination of LSTM and U-Net was able to rapidly process image time series for large areas in Cerrado. In addition, the comparison of our deforestation map with PRODES 2019 showed high agreement between them. Hence, these facts reveal the potential of our method to be applied to the entire Cerrado biome and then to automate the PRODES deforestation detection process that is currently performed by visual interpretation.

We also observed that past deforestation maps were used with success to train the algorithm. As PRODES Cerrado provides deforestation data from 2000 onward, more training samples can be selected, taking advantage of the long-term earth observation programs. For future work, we propose to divide the Cerrado biome into ecoregions [13] and apply our method for each one of these ecoregions separately in order to successfully detect the deforestation of the entire Cerrado biome using Sentinel-2 imagery.

Author Contributions: Conceptualization, B.M.M., L.M.G.F., E.C.T., R.V.M., H.d.N.B. and M.A.; methodology, B.M.M., L.M.G.F., E.C.T., R.V.M., H.d.N.B. and M.A.; software, B.M.M., E.C.T. and R.V.M.; validation, B.M.M. and M.A.; formal analysis, B.M.M., E.C.T. and H.d.N.B.; investigation, B.M.M., L.M.G.F. and M.A.; resources, B.M.M. and L.M.G.F.; data curation, B.M.M.; writing—original draft preparation, B.M.M., L.M.G.F. and M.A.; writing—review and editing, B.M.M., L.M.G.F., E.C.T., R.V.M., H.d.N.B. and M.A.; visualization, B.M.M.; supervision, L.M.G.F. and M.A.; project administration, L.M.G.F.; funding acquisition, L.M.G.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior”—Brazil (CNPq)—Grants 130574/2019-8 (B.M.M.) and 306334/2020-8 (M.A.), “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior”—Brazil (CAPES)—Finance Code 001, by the project Environmental Monitoring of the Brazilian Biomes (Amazonia Fund, BNDES—Brazilian Development Bank (No. 17.2.0536.1)) and also the project Development of systems to prevent forest fires and monitor vegetation cover in the Brazilian Cerrado (FIP—Forest Investment Program, World Bank (P143185)).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Landsat and Sentinel images are freely available from the Brazil Data Cube project portal (<https://brazildatacube.dpi.inpe.br/portal/explore>, accessed on 10 June 2021). Cerrado deforestation data freely available from PRODES portal (<http://terrabrasilis.dpi.inpe.br/en/download-2/>, accessed on 10 June 2021). SRTM terrain slope data is freely available in the Google Earth Engine platform (<https://code.earthengine.google.com/2049cabd3e13f222468f49ae37c0ad5b>, accessed on 10 June 2021). Jupyter Notebooks used to create time series stacks and use the LSTM and U-Net are available in <https://github.com/menimato/Deforestation-TimeSeries-DL/tree/db858ad41a7f266a63c20e3879181bf5940b2a92> (accessed on 10 August 2021).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

B	Blue
BDC	Brazil Data Cube
BFAST	Breaks for Additive Season and Trend
BIP	Brazilian Investment Plan
DETER	Near Real-Time Deforestation Detection
DL	Deep Learning
e	Standard error
EVI	Enhanced Vegetation Index
FIP	Forest Investment Program
FMask	Function of Mask
GIZ	German Corporation for International Cooperation
G	Green
INPE	National Institute for Space Research
JUST	Jumps Upon Spectrum and Trend
KFW	Credit Institute for Reconstruction
LSTM	Long Short-Term Memory
LULC	Land Use and Land Cover
MCTI	Ministry of Science, Technology and Innovation
MMA	Ministry of the Environment
MSI	MultiSpectral Instrument
N	Population size
n	Number of validation points

NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared
OLI	Operational Land Imager
PROBIO	Conservation and Sustainable Use of Brazilian Biological Diversity Project
PRODES	Satellite Deforestation Monitoring Project
R	Red
REDD+	Reducing Emissions from Deforestation and Forest Degradation
SE	Standard Error
SIAD	Integrated System of Deforestation Alerts
SRTM	Shuttle Radar Topography Mission
SWIR	Short Wave Infrared
tanh	Hyperbolic Tangent
UNFCCC	United Nations Framework Convention on Climate Change
z	Deviation from the mean value for the desired confidence level
α	Significance level
μ	Micro
σ^2	Variance

References

1. Strassburg, B.B.N.; Brooks, T.; Feltran-Barbieri, R.; Iribarrem, A.; Crouzeilles, R.; Loyola, R.; Latawiec, A.E.; Oliveira Filho, F.J.B.; Scaramuzza, C.A.d.M.; Scarano, F.R.; et al. Moment of Truth for the Cerrado Hotspot. *Nat. Ecol. Evol.* **2017**, *1*, 0099. [CrossRef] [PubMed]
2. Mittermeier, R.A.; Turner, W.R.; Larsen, F.W.; Brooks, T.M.; Gascon, C. Global Biodiversity Conservation: The Critical Role of Hotspots. In *Biodiversity Hotspots: Distribution and Protection of Conservation Priority Areas*; Zachos, F.E., Habel, J.C., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 3–22. [CrossRef]
3. Instituto Brasileiro de Geografia e Estatística—IBGE. Brasil em Síntese. Available online: <https://brasilemsintese.ibge.gov.br/territorio.html> (accessed on 10 June 2021).
4. Agência Nacional de Águas—ANA. Regiões Hidrográficas. Available online: http://dadosabertos.ana.gov.br/datasets/b78ea64219b9498c8125cdef390715b7_0 (accessed on 10 June 2021).
5. Ferreira, L. Seasonal landscape and spectral vegetation index dynamics in the Brazilian Cerrado: An analysis within the Large-Scale Biosphere–Atmosphere Experiment in Amazônia (LBA). *Remote Sens. Environ.* **2003**, *87*, 534–550. [CrossRef]
6. Oliveira, R.S.; Bezerra, L.; Davidson, E.A.; Pinto, F.; Klink, C.A.; Nepstad, D.C.; Moreira, A. Deep root function in soil water dynamics in cerrado savannas of central Brazil. *Funct. Ecol.* **2005**, *19*, 574–581. [CrossRef]
7. Miranda, S.d.C.d.; Bustamante, M.; Palace, M.; Hagen, S.; Keller, M.; Ferreira, L.G. Regional Variations in Biomass Distribution in Brazilian Savanna Woodland. *Biotropica* **2014**, *46*, 125–138. [CrossRef]
8. Rada, N. Assessing Brazil’s Cerrado agricultural miracle. *Food Policy* **2013**, *38*, 146–155. [CrossRef]
9. Rocha, G.F.; Guimarães Ferreira, L.; Clementino Ferreira, N.; Eduardo Ferreira, M. Detecção de Desmatamentos no Bioma Cerrado entre 2002 e 2009: Padrões, Tendências e Impactos. *Rev. Bras. Cartogr.* **2012**, *63*, 341–349.
10. Scaramuzza, C.A.d.M.; Sano, E.E.; Adami, M.; Bolfe, E.L.; Coutinho, A.C.; Esquerdo, J.C.D.M.; Maurano, L.E.P.; Narvaes, I.S.; Oliveira, F.J.B.; Rosa, R.; et al. Land-Use and Land-Cover Mapping of the Brazilian Cerrado Based Mainly on Landsat-8 Satellite Images. *Rev. Bras. Cartogr.* **2017**, *69*, 1041–1051.
11. Instituto Nacional de Pesquisas Espaciais—INPE. Monitoring Program of the Amazon and Other Biomes. Deforestation—Cerrado. Available online: http://terrabrasilis.dpi.inpe.br/download/dataset/cerrado-prodes/vector/hydrography_cerrado_biome.zip (accessed on 10 June 2021).
12. Spera, S. Agricultural Intensification Can Preserve the Brazilian Cerrado: Applying Lessons from Mato Grosso and Goiás to Brazil’s Last Agricultural Frontier. *Trop. Conserv. Sci.* **2017**, *10*, 194008291772066. [CrossRef]
13. Sano, E.E.; Rodrigues, A.A.; Martins, E.S.; Bettiol, G.M.; Bustamante, M.M.; Bezerra, A.S.; Couto, A.F.; Vasconcelos, V.; Schüler, J.; Bolfe, E.L. Cerrado ecoregions: A spatial framework to assess and prioritize Brazilian savanna environmental diversity for conservation. *J. Environ. Manag.* **2019**, *232*, 818–828. [CrossRef] [PubMed]
14. Maurano, L.E.P.; Escada, M.I.S.; Renno, C.D. Padrões espaciais de desmatamento e a estimativa da exatidão dos mapas do PRODES para Amazônia Legal Brasileira. *Ciênc. Florest.* **2019**, *29*, 1763. [CrossRef]
15. Instituto Nacional de Pesquisas Espaciais—INPE. PRODES Annual Increment of Deforested Areas in the Brazilian Cerrado. Available online: <http://www.obt.inpe.br/cerrado> (accessed on 10 June 2021).
16. Sano, E.E.; Rosa, R.; Brito, J.L.S.; Ferreira, L.G. Mapeamento semidetalhado do uso da terra do Bioma Cerrado. *Pesqui. Agropecu. Bras.* **2008**, *43*, 153–156. [CrossRef]
17. Ferreira, N.C.; Ferreira, L.G.; Huete, A.R.; Ferreira, M.E. An operational deforestation mapping system using MODIS data and spatial context analysis. *Int. J. Remote Sens.* **2007**, *28*, 47–62. [CrossRef]

18. Maurano, L.E.P.; Almeida, C.A.d.; Meira, M.B. Monitoramento do Desmatamento no Cerrado Brasileiro por Satélite—Projeto Monitoramento do Cerrado. In *Proceedings of the Simpósio Brasileiro de Sensoriamento Remoto*; INPE: São José dos Campos, Brazil, 2019; pp. 191–194.
19. Ministério do Meio Ambiente—MMA. Government Publicizes Deforestation in Cerrado. Available online: <http://redd.mma.gov.br/en/component/content/article/160-central-content/top-news/1021-government-publicizes-deforestation-in-cerrado> (accessed on 20 June 2021).
20. Ministério da Ciência, Tecnologia e Inovações—MCTI. FIP—Monitoramento Cerrado. Available online: <https://monitoramentocerrado.mcti.gov.br/> (accessed on 13 June 2021).
21. Ministério do Meio Ambiente—MMA. Desenvolvimento de Sistemas de Prevenção de Incêndios Florestais e Monitoramento da Cobertura Vegetal no Cerrado Brasileiro. Available online: <http://fip.mma.gov.br/projeto-fm/> (accessed on 13 June 2021).
22. Ministério do Meio Ambiente—MMA. Programa de Investimento Florestal no Brasil. Available online: <http://fip.mma.gov.br/> (accessed on 13 June 2021).
23. Instituto Nacional de Pesquisas Espaciais—INPE. DETER Monitoring Program of the Amazon and Other Biomes. Notices—Cerrado. Available online: <http://terrabrazilis.dpi.inpe.br/downloads/> (accessed on 10 June 2021).
24. Parente, L.; Nogueira, S.; Baumann, L.; Almeida, C.; Maurano, L.; Affonso, A.G.; Ferreira, L. Quality assessment of the PRODES Cerrado deforestation data. *Remote Sens. Appl. Soc. Environ.* **2021**, *21*, 100444. [[CrossRef](#)]
25. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep Learning in Remote Sensing Applications: A Meta-Analysis and Review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. [[CrossRef](#)]
26. Ball, J.E.; Anderson, D.T.; Chan, C.S. Comprehensive survey of deep learning in remote sensing: Theories, tools, and challenges for the community. *J. Appl. Remote Sens.* **2017**, *11*, 042609. [[CrossRef](#)]
27. Sano, E.E.; Rosa, R.; Brito, J.L.; Ferreira, L.G. Land Cover Mapping of the Tropical Savanna Region in Brazil. *Environ. Monit. Assess.* **2010**, *166*, 113–124. [[CrossRef](#)]
28. Müller, H.; Rufin, P.; Griffiths, P.; Siqueira, A.J.B.; Hostert, P. Mining Dense Landsat Time Series for Separating Cropland and Pasture in a Heterogeneous Brazilian Savanna Landscape. *Remote Sens. Environ.* **2015**, *156*, 490–499. [[CrossRef](#)]
29. Reynolds, J.; Wesson, K.; Desbiez, A.; Ochoa-Quintero, J.; Leimgruber, P. Using Remote Sensing and Random Forest to Assess the Conservation Status of Critical Cerrado Habitats in Mato Grosso do Sul, Brazil. *Land* **2016**, *5*, 12. [[CrossRef](#)]
30. Souza, C.M.; Shimbo, J.Z.; Rosa, M.R.; Parente, L.L.; Alencar, A.A.; Rudorff, B.F.T.; Hasenack, H.; Matsumoto, M.; Ferreira, L.G.; Souza-Filho, P.W.M.; et al. Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. *Remote Sens.* **2020**, *12*, 2735. [[CrossRef](#)]
31. Parente, L.; Taquary, E.; Silva, A.; Souza, C.; Ferreira, L. Next Generation Mapping: Combining Deep Learning, Cloud Computing, and Big Remote Sensing Data. *Remote Sens.* **2019**, *11*, 2881. [[CrossRef](#)]
32. Belward, A.S.; Skøien, J.O. Who launched what, when and why: Trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 115–128. [[CrossRef](#)]
33. Maretto, R.V.; Fonseca, L.M.G.; Jacobs, N.; Korting, T.S.; Bendini, H.N.; Parente, L.L. Spatio-Temporal Deep Learning Approach to Map Deforestation in Amazon Rainforest. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 771–775. [[CrossRef](#)]
34. Bendini, H.N.; Fonseca, L.M.; Schwieder, M.; Rufin, P.; Korting, T.S.; Koumrouyan, A.; Hostert, P. Combining environmental and landsat analysis ready data for vegetation mapping: A case study in the Brazilian Savanna Biome. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2020**, *43*, 953–960. [[CrossRef](#)]
35. Zeng, L.; Wardlow, B.D.; Xiang, D.; Hu, S.; Li, D. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sens. Environ.* **2020**, *237*, 111511. [[CrossRef](#)]
36. Verbesselt, J.; Hyndman, R.; Newnham, G.; Culvenor, D. Detecting trend and seasonal changes in satellite image time series. *Remote Sens. Environ.* **2010**, *114*, 106–115. [[CrossRef](#)]
37. Masiliūnas, D.; Tsendbazar, N.E.; Herold, M.; Verbesselt, J. BFAST Lite: A Lightweight Break Detection Method for Time Series Analysis. *Remote Sens.* **2021**, *13*, 3308. [[CrossRef](#)]
38. Ghaderpour, E.; Vujadinovic, T. Change Detection within Remotely Sensed Satellite Image Time Series via Spectral Analysis. *Remote Sens.* **2020**, *12*, 4001. [[CrossRef](#)]
39. de Bem, P.; de Carvalho, O., Jr.; Guimarães, R.F.; Gomes, R.T. Change Detection of Deforestation in the Brazilian Amazon Using Landsat Data and Convolutional Neural Networks. *Remote Sens.* **2020**, *12*, 901. [[CrossRef](#)]
40. Adarme, M.O.; Feitosa, R.Q.; Happ, P.N.; Almeida, C.A.D.; Gomes, A.R. Evaluation of Deep Learning Techniques for Deforestation Detection in the Brazilian Amazon and Cerrado Biomes From Remote Sensing Imagery. *Remote Sens.* **2020**, *12*, 910. [[CrossRef](#)]
41. Rußwurm, M.; Körner, M. Self-attention for raw optical Satellite Time Series Classification. *ISPRS J. Photogramm. Remote Sens.* **2020**, *169*, 421–435. [[CrossRef](#)]
42. Interdonato, R.; Ienco, D.; Gaetano, R.; Ose, K. DuPLO: A DuAl view Point deep Learning architecture for time series classification. *ISPRS J. Photogramm. Remote Sens.* **2019**, *149*, 91–104. [[CrossRef](#)]
43. Xu, J.; Zhu, Y.; Zhong, R.; Lin, Z.; Xu, J.; Jiang, H.; Huang, J.; Li, H.; Lin, T. DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping. *Remote Sens. Environ.* **2020**, *247*, 111946. [[CrossRef](#)]
44. Dutta, D.; Chen, G.; Chen, C.; Gagné, S.A.; Li, C.; Rogers, C.; Matthews, C. Detecting Plant Invasion in Urban Parks with Aerial Image Time Series and Residual Neural Network. *Remote Sens.* **2020**, *12*, 3493. [[CrossRef](#)]

45. Taquary, E.C. Deep Learning para Identificação Precisa de Desmatamentos Através do Uso de Imagens Satelitárias de Alta Resolução. Master's Thesis, Universidade Federal de Goiás (UFG), Goiânia, Brazil, 2019.
46. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
47. Graves, A.; Rahman Mohamed, A.; Hinton, G. Speech recognition with deep recurrent neural networks. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013. [[CrossRef](#)]
48. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241. [[CrossRef](#)]
49. National Aeronautics and Space Administration—NASA. Landsat 8. Available online: <https://landsat.gsfc.nasa.gov/landsat-8> (accessed on 10 July 2021).
50. European Space Agency—ESA. Sentinel-2 MSI Introduction. Available online: <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi> (accessed on 10 July 2021).
51. MapBiomias. Collection 4.0 of the Annual Series of Land Use and Land Cover in Brazil. Available online: <http://plataforma.mapbiomas.org> (accessed on 10 June 2021).
52. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.C.; Skakun, S.V.; Justice, C. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* **2018**, *219*, 145–161. [[CrossRef](#)]
53. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.Y.; Wong, W.K.; Woo, W.C. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; pp. 802–810.
54. Martinez, J.A.C.; Rosa, L.E.C.L.; Feitosa, R.Q.; Sanches, I.D.; Happ, P.N. Fully convolutional recurrent networks for multitemporal crop recognition from multitemporal image sequences. *ISPRS J. Photogramm. Remote Sens.* **2021**, *171*, 188–201. [[CrossRef](#)]
55. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, RG2004. [[CrossRef](#)]
56. Ferreira, K.R.; Queiroz, G.R.; Vinhas, L.; Marujo, R.F.B.; Simoes, R.E.O.; Picoli, M.C.A.; Camara, G.; Cartaxo, R.; Gomes, V.C.F.; Santos, L.A.; et al. Earth Observation Data Cubes for Brazil: Requirements, Methodology and Products. *Remote Sens.* **2020**, *12*, 4033. [[CrossRef](#)]
57. Qiu, S.; Zhu, Z.; He, B. Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery. *Remote Sens. Environ.* **2019**, *231*, 111205. [[CrossRef](#)]
58. Dozier, J.; Painter, T.H.; Rittger, K.; Frew, J.E. Time–space continuity of daily maps of fractional snow cover and albedo from MODIS. *Adv. Water Resour.* **2008**, *31*, 1515–1526. [[CrossRef](#)]
59. Hou, J.; Huang, C.; Zhang, Y.; Guo, J.; Gu, J. Gap-Filling of MODIS Fractional Snow Cover Products via Non-Local Spatio-Temporal Filtering Based on Machine Learning Techniques. *Remote Sens.* **2019**, *11*, 90. [[CrossRef](#)]
60. Zhang, H.K.; Roy, D.P.; Yan, L.; Li, Z.; Huang, H.; Vermote, E.; Skakun, S.; Roger, J.C. Characterization of Sentinel-2A and Landsat-8 top of atmosphere, surface, and nadir BRDF adjusted reflectance and NDVI differences. *Remote Sens. Environ.* **2018**, *215*, 482–494. [[CrossRef](#)]
61. Berlinck, C.N.; Batista, E.K. Good fire, bad fire: It depends on who burns. *Flora* **2020**, *268*, 151610. [[CrossRef](#)]
62. França, H.; Setzer, A.W. AVHRR analysis of a savanna site through a fire season in Brazil. *Int. J. Remote Sens.* **2001**, *22*, 2449–2461. [[CrossRef](#)]
63. Bittencourt, O.O.; Morelli, F.; Júnior, C.A.S.; Santos, R. An Approach to Classify Burned Areas Using Few Previously Validated Samples. In *Computational Science and Its Applications—ICCSA 2020*; Springer International Publishing: Cham, Switzerland, 2020; pp. 239–254. [[CrossRef](#)]
64. Pereira, A.; Pereira, J.; Libonati, R.; Oom, D.; Setzer, A.; Morelli, F.; Machado-Silva, F.; de Carvalho, L. Burned Area Mapping in the Brazilian Savanna Using a One-Class Support Vector Machine Trained by Active Fires. *Remote Sens.* **2017**, *9*, 1161. [[CrossRef](#)]
65. Li, Y.; Zhang, Y.; Zhu, Z. Error-Tolerant Deep Learning for Remote Sensing Image Scene Classification. *IEEE Trans. Cybern.* **2021**, *51*, 1756–1768. [[CrossRef](#)] [[PubMed](#)]
66. Rendón, E.; Alejo, R.; Castorena, C.; Isidro-Ortega, F.J.; Granda-Gutiérrez, E.E. Data Sampling Methods to Deal with the Big Data Multi-Class Imbalance Problem. *Appl. Sci.* **2020**, *10*, 1276. [[CrossRef](#)]
67. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. 2015. Available online: <https://www.tensorflow.org> (accessed on 10 June 2021).
68. Ribeiro, J.F.; Walter, B.M.T. As Principais Fitofisionomias do Bioma Cerrado. In *Cerrado: Ecologia e Flora*; Sano, S.M., Almeida, S.P., Ribeiro, J.F., Eds.; EMBRAPA: Brasília, Brazil, 2008; pp. 152–212.
69. Maretto, R.V.; Korting, T.S.; Fonseca, L.M.G. An Extensible and Easy-to-use Toolbox for Deep Learning Based Analysis of Remote Sensing Images. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2019), Yokohama, Japan, 28 July–2 August 2019. [[CrossRef](#)]
70. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [[CrossRef](#)]
71. Lohr, S.L. *Sampling: Design and Analysis*, 2nd ed.; Brooks/Cole: Boston, MA, USA, 2009; Volume 1, p. 596.

72. Alencar, A.; Shimbo, J.Z.; Lenti, F.; Marques, C.B.; Zimbres, B.; Rosa, M.; Arruda, V.; Castro, I.; Ribeiro, J.F.M.; Varela, V.; et al. Mapping Three Decades of Changes in the Brazilian Savanna Native Vegetation Using Landsat Data Processed in the Google Earth Engine Platform. *Remote Sens.* **2020**, *12*, 924. [[CrossRef](#)]
73. Jiang, L.; Zhou, Z.; Leung, T.; Li, L.J.; Fei-Fei, L. MentorNet: Learning Data-Driven Curriculum for Very Deep Neural Networks on Corrupted Labels. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018.
74. Lima, T.A.; Beuchle, R.; Langner, A.; Grecchi, R.C.; Griess, V.C.; Achard, F. Comparing Sentinel-2 MSI and Landsat 8 OLI Imagery for Monitoring Selective Logging in the Brazilian Amazon. *Remote Sens.* **2019**, *11*, 961. [[CrossRef](#)]
75. Bueno, I.; Acerbi, F., Jr.; Silveira, E.; Mello, J.; Carvalho, L.; Gomide, L.; Withey, K.; Scolforo, J. Object-Based Change Detection in the Cerrado Biome Using Landsat Time Series. *Remote Sens.* **2019**, *11*, 24–31. [[CrossRef](#)]
76. Schwieder, M.; Leitão, P.J.; da Cunha Bustamante, M.M.; Ferreira, L.G.; Rabe, A.; Hostert, P. Mapping Brazilian savanna vegetation gradients with Landsat time series. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 361–370. [[CrossRef](#)]
77. Oliveira, L.M.T.; França, G.B.; Nicácio, R.M.; Antunes, M.A.H.; Costa, T.C.C.; Torres, A.R.; França, J.R.A. A study of the El Niño-Southern Oscillation influence on vegetation indices in Brazil using time series analysis from 1995 to 1999. *Int. J. Remote Sens.* **2010**, *31*, 423–437. [[CrossRef](#)]
78. Liu, P.; Zhang, H.; Eom, K.B. Active Deep Learning for Classification of Hyperspectral Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 712–724. [[CrossRef](#)]