





Article

Evaluation of Deep Learning Techniques for Deforestation Detection in the Brazilian Amazon and Cerrado Biomes From Remote Sensing Imagery

Mabel Ortega Adarme ^{1,*} , Raul Queiroz Feitosa ¹ , Patrick Nigri Happ ¹ ,
Claudio Aparecido De Almeida ²  and Alessandra Rodrigues Gomes ²

¹ Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, 22451-900 Rio de Janeiro, Brazil; raul@ele.puc-rio.br (R.Q.F.); patrick@ele.puc-rio.br (P.N.H.)

² National Institute for Space Research (INPE), São Jose dos Campos 12227-010, São Paulo, Brazil; claudio.almeida@inpe.br (C.A.D.A.); alessandra.gomes@inpe.br (A.R.G.)

* Correspondence: mortega@ele.puc-rio.br

Received: 05 February 2020; Accepted: 6 March 2020; Published: 12 March 2020

Abstract: Deforestation is one of the major threats to natural ecosystems. This process has a substantial contribution to climate change and biodiversity reduction. Therefore, the monitoring and early detection of deforestation is an essential process for preservation. Techniques based on satellite images are among the most attractive options for this application. However, many approaches involve some human intervention or are dependent on a manually selected threshold to identify regions that suffer deforestation. Motivated by this scenario, the present work evaluates Deep Learning-based strategies for automatic deforestation detection, namely, Early Fusion (EF), Siamese Network (SN), and Convolutional Support Vector Machine (CSVM) as well as Support Vector Machine (SVM), used as the baseline. The target areas are two regions with different deforestation patterns: the Amazon and Cerrado biomes in Brazil. The experiments used two co-registered Landsat 8 images acquired at different dates. The strategies based on Deep Learning achieved the best performance in our analysis in comparison with the baseline, with SN and EF superior to CSVM and SVM. In the same way, a reduction of the salt-and-pepper effect in the generated probabilistic change maps was noticed as the number of training samples increased. Finally, the work assesses how the methods can reduce the time invested in the visual inspection of deforested areas.

Keywords: deforestation detection; Brazilian biomes; deep learning; optical imagery

1. Introduction

Deforestation is one of the largest sources of anthropogenic CO₂ emissions. It is a wide-reaching problem, including the reduction of carbon storage, greenhouse gas emissions, and other environmental issues such as biodiversity losses [1]. Currently, one of the highest deforestation rates occurs in South America [2], where the most significant statistics of tree losses are concentrated in Brazil [3,4]. This country comprises most of the Amazon rainforest, with 60% of its total territory [5]. In particular, Amazon and Cerrado biomes cover the most significant portion of the Brazilian territory, with an area of about 49% and 24%, respectively, comprising together an area of around 6.2 million square kilometers of the Brazilian territory. In both biomes, the deforested areas are predominantly converted to pasture [6,7], in addition to a strong expansion of soy in the Cerrado biome [8]. These biomes have different characteristics and accommodate rich biodiversity of endemic species, many of them vulnerable [9]. Therefore, the conservation of Amazon and Cerrado biomes is essential for the future of our planet.

For several decades, the Amazon ecosystems have been threatened by disorderly economic growth. The main causes are the growth of agribusiness, certain mining and logging activities,

forest fires, the emergence of informal settlements, and the expansion of physical infrastructure to meet population growth [10–12]. According to the National Institute for Space Research (INPE) [13], deforestation increased considerably in the period of the 1990s and early 2000s in the Amazon biome, but it declined steadily from 2004 to 2014. However, the World Wildlife Fund (WWF) [14] estimates that about a quarter of the biome will be disappearing by 2030 if the actual rate of deforestation keeps increasing.

Following the Amazon biome, Cerrado is the Brazilian biome that undergone significant changes due to human occupation [15]. Despite its biological importance, it faces intense land-use pressure, losing over 50% of natural vegetation due to agricultural expansion [16–18]. Moreover, the Cerrado has been one of the least studied biomes in Brazil. It has not received proper attention compared to other Brazilian biomes, such as the Amazon and the Atlantic Forest [9,17]. Information about land use and land cover changes of this biome are still limited. In this sense, the monitoring of this biome is indispensable to track natural disasters and human activities, and thus, supporting the development and implementation of government policies to prevent illegal activities [19]. In this scenario, Remote Sensing (RS) is an essential source of data for the effective monitoring of these regions.

Some of the traditional unsupervised techniques have been widely used to detect changes in multi-temporal images. These techniques include Image Differencing [20], Image Rationing [21], Regression Analysis [22], Change Vector Analysis (CVA) [23], and Principal Component Analysis (PCA) [24,25]. However, they strongly depend on manually selected thresholds to define what is considered a change [26] and the engineered features these algorithms are based upon generally result in a weak description of the objects represented in the image [27].

For the supervised methods, Support Vector Machines (SVMs) have been often used for the classification of satellite images [28,29] mainly because of its high accuracy in scenarios where there are few available labeled samples. Nevertheless, a proper setup of its hyperparameters, including the appropriate kernel function, is essential to achieve good generalization. In general, this process involves considering several kernel types and comparing them via cross-validation or other methods [30]. Similarly, tree-based [31] classifiers and methods based on Artificial Neural Networks (ANN) are widely used for image classification [32].

On the other hand, several governmental projects carry out systematic satellite monitoring of deforestation in the Amazon and Cerrado Biomes. In particular, the Amazon Deforestation Monitoring Project (PRODES (<http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>)) provides annual reports about clear cut deforestation in the Brazilian Legal Amazon (BLA) since 1988 [33]. The deforested areas are estimated by photointerpretation of images acquired at different dates. This process is carried out by specialists who delineate polygons of deforestation directly on the computer screen [16], which makes the whole task highly dependent on human intervention. The project currently uses LANDSAT 8/OLI, SENTINEL 2 and CBERS 4 images [34] and it considers a minimum mapping area of 6.25 ha, regardless of the used instrument. The results are considered reliable [35] and recent analyzes indicate an accuracy level close to 95% [13].

Another related project is the Near Real Time Deforestation Detection Project (DETER) [36], which main objective is to reinforce the effective land-use policies in BLA. It supervises the irregular deforestation, vegetation deterioration, and wood exploration. This process is done by using Linear Spectral Mixture Model (LSMM) and visual interpretation based on five main elements (color, tonality, texture, shape, and context). Finally, the Land Use and Land Cover Mapping of Amazon Deforested Areas (TerraClass) project identifies deforestation in BLA and searches for potential reasons for logging. De Almeida et al. [6], showed that in the Amazonian post-deforestation landscape, pasture is dominant, followed by Secondary Vegetation. Eventually, part of them regenerates and turns into the forest again.

Similar to BLA, the Cerrado biome has also been monitored by INPE through the PRODES Cerrado (<http://www.obt.inpe.br/cerrado>). Brito et al. [37] presented the methodology used for mapping deforestation in the Cerrado biome since 2000. The images adopted in the project are from Landsat-5/TM, Landsat-7/ETM+, Landsat-8/OLI and Resourcesat-2/LISS-III, and a minimum

mapping area of 1 ha is considered. The process is entirely manual and is carried out by visual interpretation by taking into account five main elements: color, tonality, texture, shape, and context. TerraClass Cerrado is another project coordinated by Brazil's environment ministry in cooperation with the Brazilian Agricultural Research Corporation and INPE. With this project, they map the use and cover of deforested areas of the Brazilian Amazon, and they enable the characterization of the areas mapped by PRODES, using satellite images from 2013 (<http://www.dpi.inpe.br/tccerrado/>).

MapBiomias (<http://mapbiomas.org/>) is another initiative that analyzes the Brazilian territory by mapping the land use and land cover since 1985. The methodology adopted by MapBiomias is presented in [38]. Its dataset comprises images of all Brazilian biomes collected by Landsat 5-6-7 sensors. The methodology also involves extracting statistical features to train a Random Forest (RF) classifier. In a post-classification stage, spatial and temporal filters remove classification noise and fill information gaps due to clouds. In this methodology, all procedures are performed using the Google Earth Engine. However, the final validation stage is based on visual interpretation, and an overestimated forest area is presented.

Further works have been developed by the RS community. Bueno et al. [9] presented a study for detecting deforestation in areas of the Cerrado biome. They adopted the object-based image analysis (OBIA) methodology, applied to Landsat OLI (The Operational Land Imager) time series, to find out the best spectral bands and vegetation indices for discrimination of true deforestation from seasonal changes using a RF classifier. Likewise, Machado et al. [39] presented a study of mapping the deforested areas using images of MODIS sensor (or Moderate Resolution Imaging Spectroradiometer). A maximum likelihood classifier implemented in ERDAS software was used for the task. Moreover, Picoli et al. [40] presented a land use and land cover classification using satellite image time series. They used a SVM model to discriminate natural and human-transformed land areas in the state of Mato Grosso.

More recently, Deep Learning (DL) techniques have become state-of-the-art in many application fields, including RS. Through the potential of Deep Neural Networks (DNNs), representations at multiple levels can be extracted, which usually provides features with further information and often allows for better results than what can be achieved by using domain-specific handcrafted features. Zagoruyko and Komodakis [41], introduced multiple CNN models to learn similar patterns in pairs of images, which experienced geometric transformation and changes in their illumination patterns, among other changes. The authors reported promising results in comparison to traditional methods that rely on hand-crafted features. Some of those methods are the Early Fusion, Siamese CNN, and Pseudo-Siamese, which were later used by Caye et al. [42] for urban change detection. In this work, the authors compared the Siamese CNN and Early Fusion CNN techniques and evaluated the impact of using different spectral channels as inputs.

In a similar work [43], a Siamese CNN was successfully applied to detect changes in objects such as buildings and trees, as well as to discriminate real and false changes generated by inaccurate registrations or alignments. To this goal, the patches assigned as "change" were grouped and verified as individual object changes.

Zhan et al. [27] proposed a supervised change detection method based on a deep Siamese CNN for optical aerial images. As in previous approaches, the authors of this work improved the preliminary classification results in a post-processing stage. Essentially, the score map produced by the SN is segmented using a thresholding technique. The generated segments are then classified using a k -nearest neighbor (k -NN) approach. In a similar approach, the authors of [44] integrate the advantages of CNN and RNN to learn joint spectral-spatial-temporal features and solve a multispectral image change detection problem, achieving encouraging results.

Goals And Contributions

This work evaluates Deep-Learning based techniques applied to deforestation detection in two tropical regions with different deforestation patterns: the Amazon and Cerrado biomes.

For comparison purposes, we used a SVM classifier, which was taken as the baseline. The main objective is to reduce the human effort involved in monitoring programs such as the Amazon Deforestation Monitoring Project and the Cerrado Monitoring Project. In addition, this work aims at contributing to improve accuracy and reduce the subjectivity inherent to human photointerpretation, besides the reduction of costs and time for monitoring the vegetation of these biomes, which meets the need for constant improvement of monitoring instruments [45].

The main contributions of this work are:

- An evaluation and comparison of three Deep Learning techniques for automatic deforestation detection in Brazilian Amazon and Cerrado biomes; namely, Early Fusion (EF), Siamese Network (SN), and Convolutional SVM (CSVM).
- An assessment of these methods' accuracy under scarce training samples.
- An estimation for each method of the relation: area assigned as deforestation vs. area of true deforestation.

The rest of the paper is structured as follows. Section 2 describes the assessed methods for deforestation detection, the study areas, and the experimental protocol. Next, the experimental results are presented and discussed in Section 3. Finally, conclusions and future works are summarized in Section 4.

2. Materials and Methods

This section explains the three DL based approaches investigated in the present analysis for detecting deforestation from optical images: Early Fusion (EF), Siamese Network (SN), and Convolutional SVM (CSVM).

For all methods, the inputs are pairs of co-registered images of two optical images acquired at different dates, denoted T1 and T2 henceforth. The classification receives as input an image patch, and the result is assigned to the patch central pixel. A sliding window approach is adopted to classify all pixels of the target site.

2.1. Early Fusion (EF)

This model can be regarded as an extension of a regular CNN. It is composed of a series of convolutions and pooling layers, followed by fully connected (FC) layers, whereby the last one is a Softmax layer having as many outputs as the number of classes. Softmax assigns posterior probability values to each class in a classification problem, which adds up to 1. For deforestation detection, this layer has two outputs related to the “deforestation” and “no-deforestation” classes, and the final label is defined based on the class corresponding to the maximum probability.

The EF architecture used in this work was inspired by the CNN model proposed in [42] used for detecting changes in urban areas with good reported performance. The architecture of this CNN model takes as input the stack formed by the concatenation of both images (T1 and T2) along the spectral dimension. Each patch is a tensor of size h -by- h -by- $2c$, denoting the patch height, width, and depth, respectively. Figure 1 outlines the procedure.

2.2. Siamese Network (SN)

A Siamese Network can be regarded as an extension of a conventional CNN. The particular network design used in this analysis was adapted from [42]. It consists of two subnetworks that share the same hyperparameters and weights [43]. Each patch of a pair of corresponding patches feeds a subnetwork. In consequence, the descriptor vectors of the two corresponding patches are computed by the same model [27].

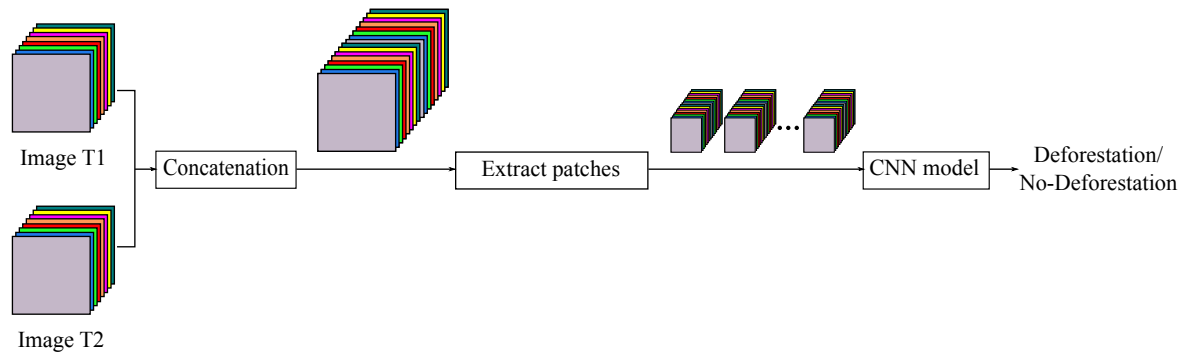


Figure 1. EF method. Co-registered images acquired at T1 and T2 are stacked along the spectral dimension; correspondent patches of both images are the input for the CNN model.

The descriptor vectors delivered by each subnetwork are concatenated to produce the final feature vector, which is forwarded to a two-layer decision network [46] that assigns the label “deforestation” or “no-deforestation” to the central pixel of the input patch pair. The whole procedure is illustrated in Figure 2.

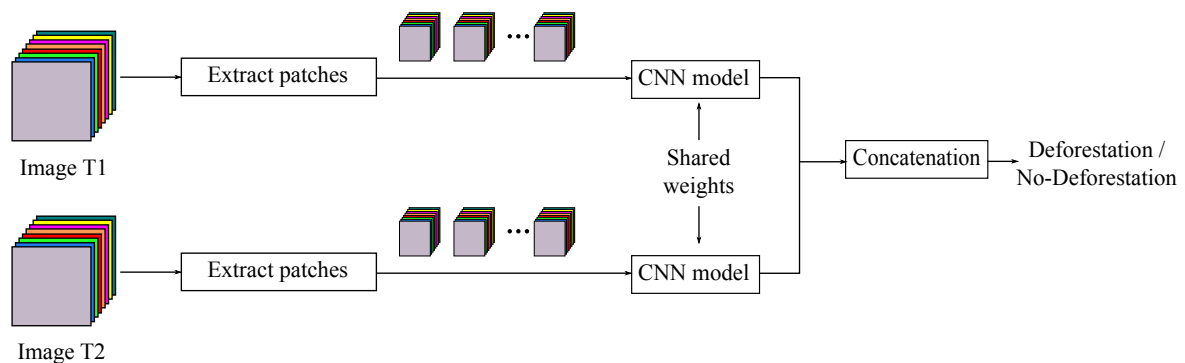


Figure 2. SN method. Correspondent patches cropped from two co-registered images acquired at T1 and T2 are the inputs for two CNN networks, both sharing the same architecture and parameter values.

2.3. Convolutional SVM (CSVM)

The CSVM, proposed by [47], is an alternative DL approach based on SVMs. This method was tested for object detection from Unmanned Aerial Vehicle (UAV) imagery and performed well in the task of discerning between instances of the object of interest and the background. Analogous to a traditional CNN, a CSVM architecture is composed of a set of convolutional and pooling layers followed by a classification layer at the end [47], but in contrast to CNN, CSVM does not use the backpropagation algorithm during the training; it trains the set of linear SVMs in a layerwise fashion. The intended advantage of this method is its performance in classification tasks where there are very few training samples available.

Similar to the EF method, the two input images (T1 and T2) are concatenated along their spectral dimension. Again, as in EF, in the CSVM approach, we classify patches of size h -by- h -by- $2c$ whose classification output is assigned to the patch central pixel. Next, we describe how the method proposed by Bazi and Melgani [47] for image classification was adapted in this work for pixel-wise deforestation detection.

2.3.1. Construction of Training Set

Following patches extraction, the training set is created for learning the SVMs filters. The extracted input patches are split into non overlapping rectangular sections, called mini-patches, of size h_1 -by- h_1 -by- $2c$, which are vectorized to form the global training set. This procedure is illustrated in Figure 3a.

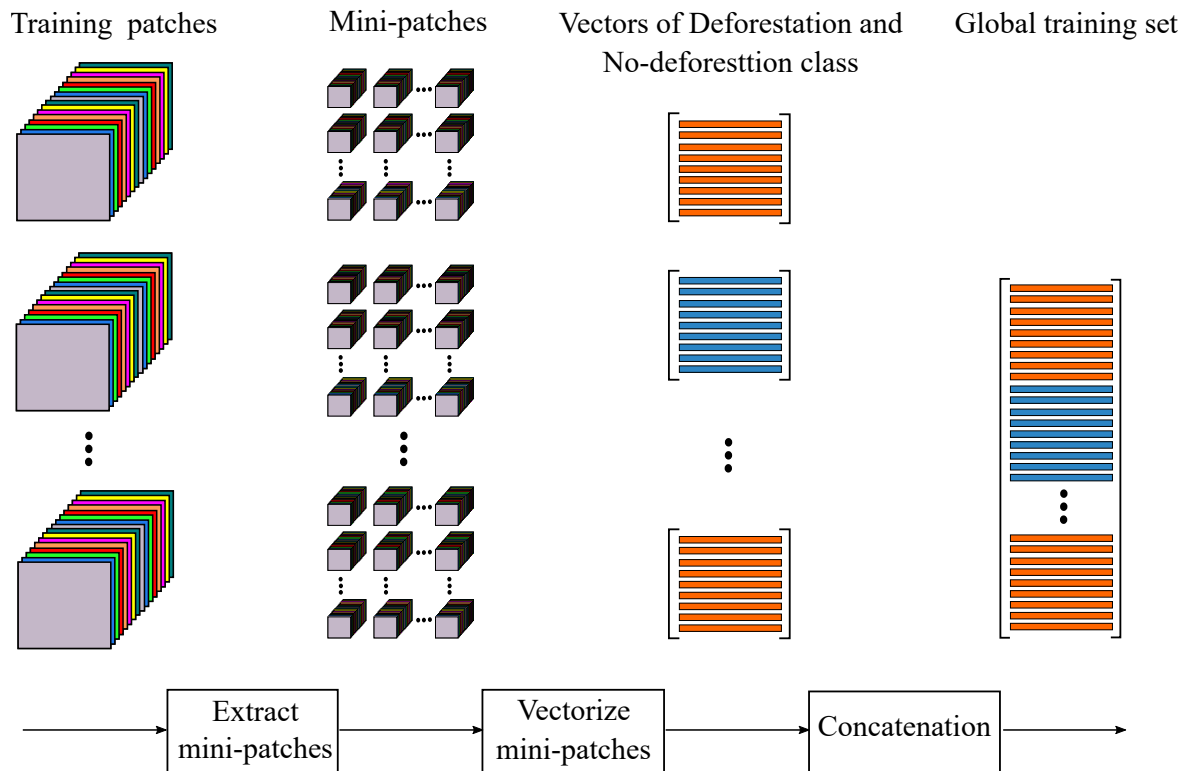


Figure 3. Procedure to train the SVMs filter bank. Mini-patches are extracted from input patches, and they are vectorized to compose the training set of SVMs.

2.3.2. Training the SVMs Filter Bank

After the global training set is built, m subsets of N random selected samples are created to train m SVMs filters. These m subsets are composed of n samples per class, which are randomly selected from the global training set. The weights of the SVMs filters are learned using a conventional forward supervised learning layer by layer in a greedy fashion. To make the most of available training samples and to avoid data duplication in the subsets, in our study, the value of n was set to the ratio between the number of training samples (N) and the amount (m) of SVM filters.

2.3.3. Generation of Feature Maps

In this stage, the input patch pairs are convolved with the learned SVM filters to generate the feature maps, which are fed to a pooling layer followed by a non-linear activation function. The output is the input to the next convolution layer (see Figure 4). The procedure is repeated until the desired number of layers is reached.

2.3.4. Classification

As mentioned before, the feature maps obtained in the last layer are fed again to a final binary SVM classifier that identifies the class label of each patch central pixel, either as “deforestation” or “no-deforestation”. In the original CSVM, the final feature map is a vector containing the means of the four quadrants of the input feature map. In contrast, in our approach, the feature descriptor is a vector obtained after flattening the output of each convolutional layer. This procedure was carried out after experimental analysis where the results presented a better performance as well as a reduction in the inference time.

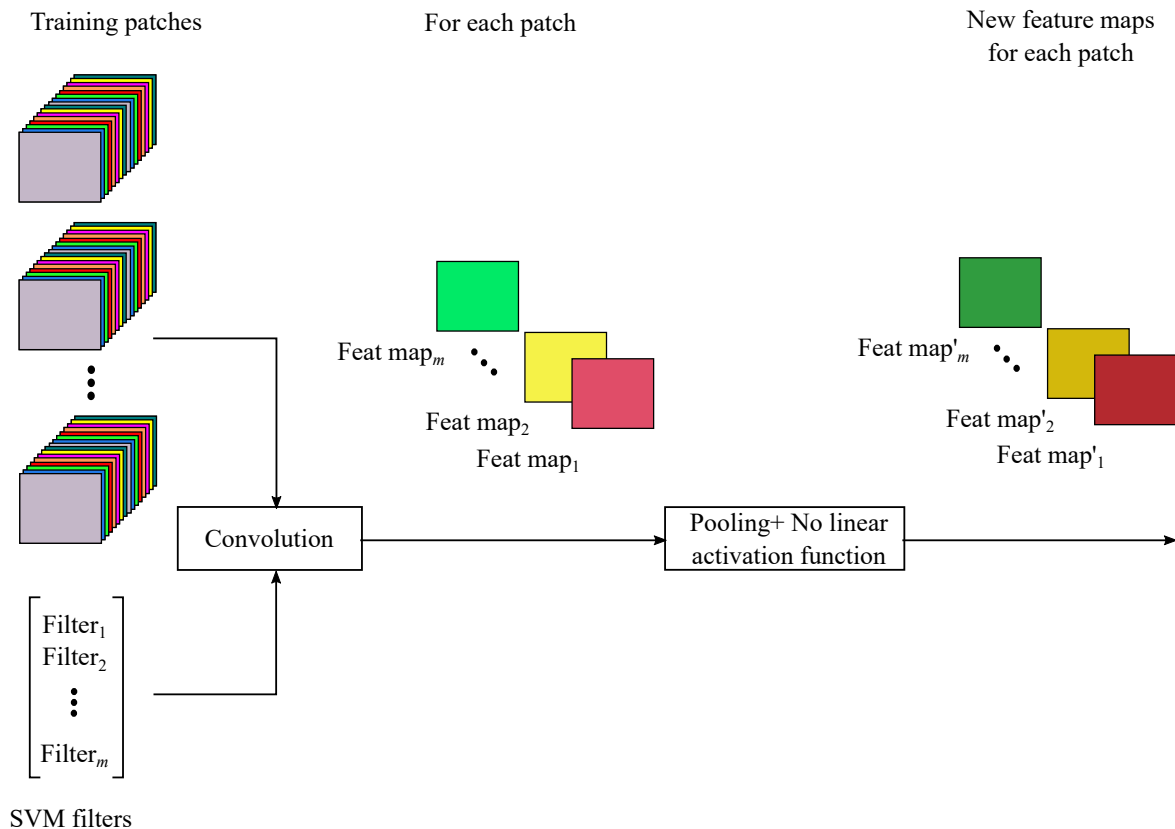


Figure 4. CSVM - Generation of feature maps.

2.4. Study Areas

Two study areas from the Brazilian Biomes were selected. The first one is a region located in the Amazon biome, and the second one is located in the Cerrado biome. The detailed description of each one is presented in the following.

2.4.1. Amazon Biome

The first study area corresponds to a region of the Amazon Biome, more specifically localized in the Pará State, Brazil, centered on coordinates of 03°17'23" South and 050°55'08" West, Figure 5. Pará state comprises 26% of Brazilian Amazon [48], and most of it is covered with dense tropical rainforest. This area has faced continuous degradation process, as indicated by PRODES and DETER reports [33].

The reference change map used in our experiments refers to the deforestation that occurred between August 2016 and July 2017. This information was downloaded from INPE site, which is freely available at the PRODES database (Available at <http://terrabrasilis.dpi.inpe.br/map/deforestation>). For this reference, the following considerations were taken into account:

- Polygons of areas deforested in previous years (before August 2016) were disregarded.
- An external buffer of two pixels inside the polygons of class "deforestation" was not considered for the training, validation, and test. The reason was to avoid the impact of the variation between the photointerpreters estimation.
- Areas lower than 6.25 ha (69 pixels) were also not considered in our evaluation because PRODES data does not record deforestation areas smaller than that for the Amazon biome.

The dataset is composed of two Landsat 8-OLI scenes, with a spatial resolution equal to 30 m. The images were acquired by the United States Geological Survey (USGS). After the atmospheric correction, the images were clipped around the selected area. The resulting data of the BLA had

a size of 1100×2600 pixels and seven spectral bands: Coastal/Aerosol, Blue, Green, Red, NIR, SWIR-1, and SWIR-2. The acquisition dates were August 2nd, 2016, and July 20th, 2017 (see Figure 5). They were selected based on PRODES reference date, which computes the annual deforestation rate from August 1st of each year, during the dry season (June to September), when the cloud cover, a major problem over the whole BLA region, is minimum.

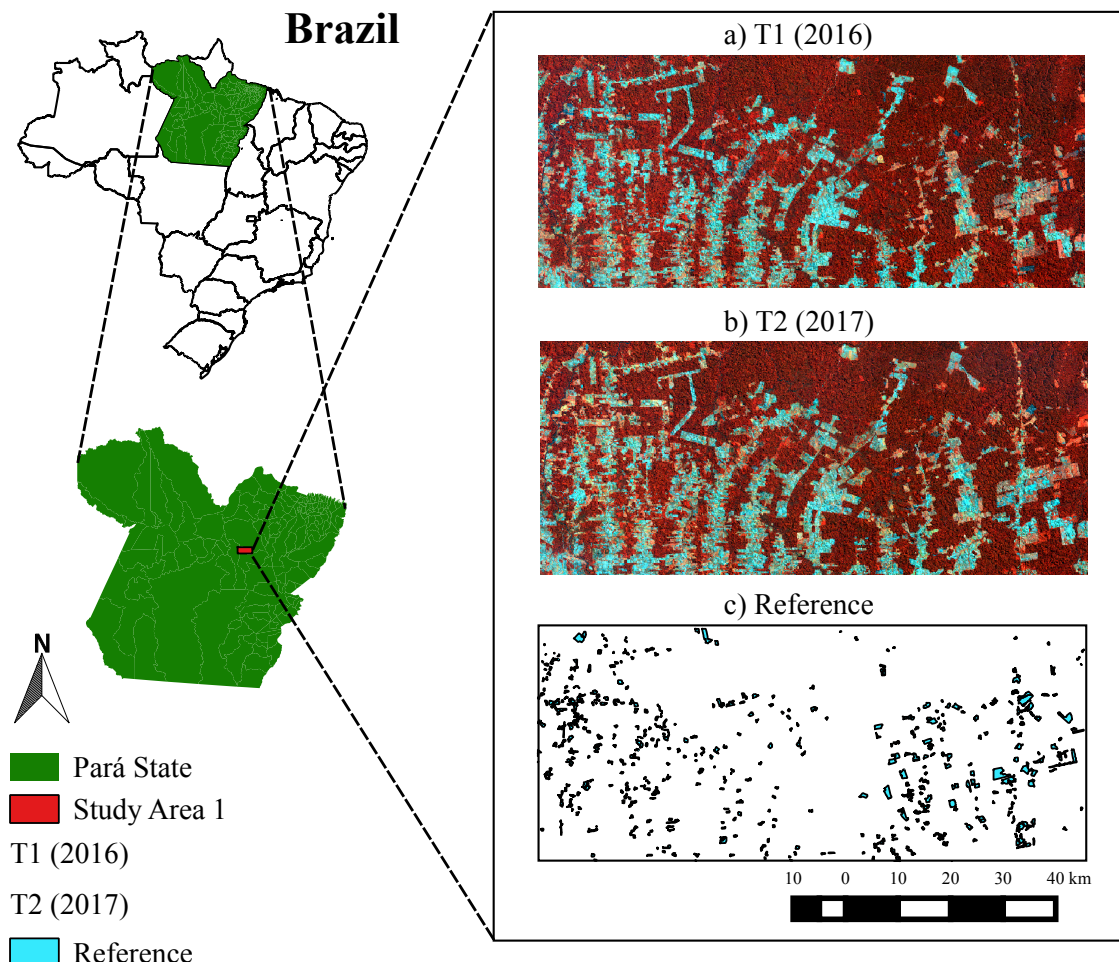


Figure 5. Region of the Brazilian Amazon biome, located in Pará State, Brazil. *NIR – G – B* composition of the study area at two different dates T1 and T2 (a,b); (c) Reference of the deforestation process from 2016 to 2017.

2.4.2. Cerrado Biome

The second study area belongs to the Brazilian Cerrado biome, localized in the Maranhão State, Brazil, centered on coordinates of $04^{\circ} 58' 53''$ S and $043^{\circ} 49' 41''$ W. Figure 6 illustrates this study area. The state of Maranhão is in a transition area among three different biomes: Cerrado (64%), Amazon (35%) and Caatinga (1%), with a predominance of savanna formations in the Cerrado. This transition makes the Cerrado Maranhense present from dense tree formations, known as “Cerradão”, and more open formations with low shrubs, vegetation with twisted trunks and thick barks typical of a “Stricto Senso” savanna. This Cerrado vegetation has suffered a significant agricultural expansion, most of it over native vegetation [49], and the deforestation in this biome has also been monitored by PRODES (Available at <http://www.obt.inpe.br/cerrado/>). The dataset is also composed of two images from Landsat 8-OLI with seven spectral bands, pre-processed in the same way as in the Amazon Biome

dataset. The size of the images was 1719×1442 pixels. For this database, the first image is from 3 September 2017, and the second one is from 22 September 2018. Since the reference provided by PRODES is also from the dry season, the reference used in this case does not contain all the deforested areas. Then, the reference had the following adaptations:

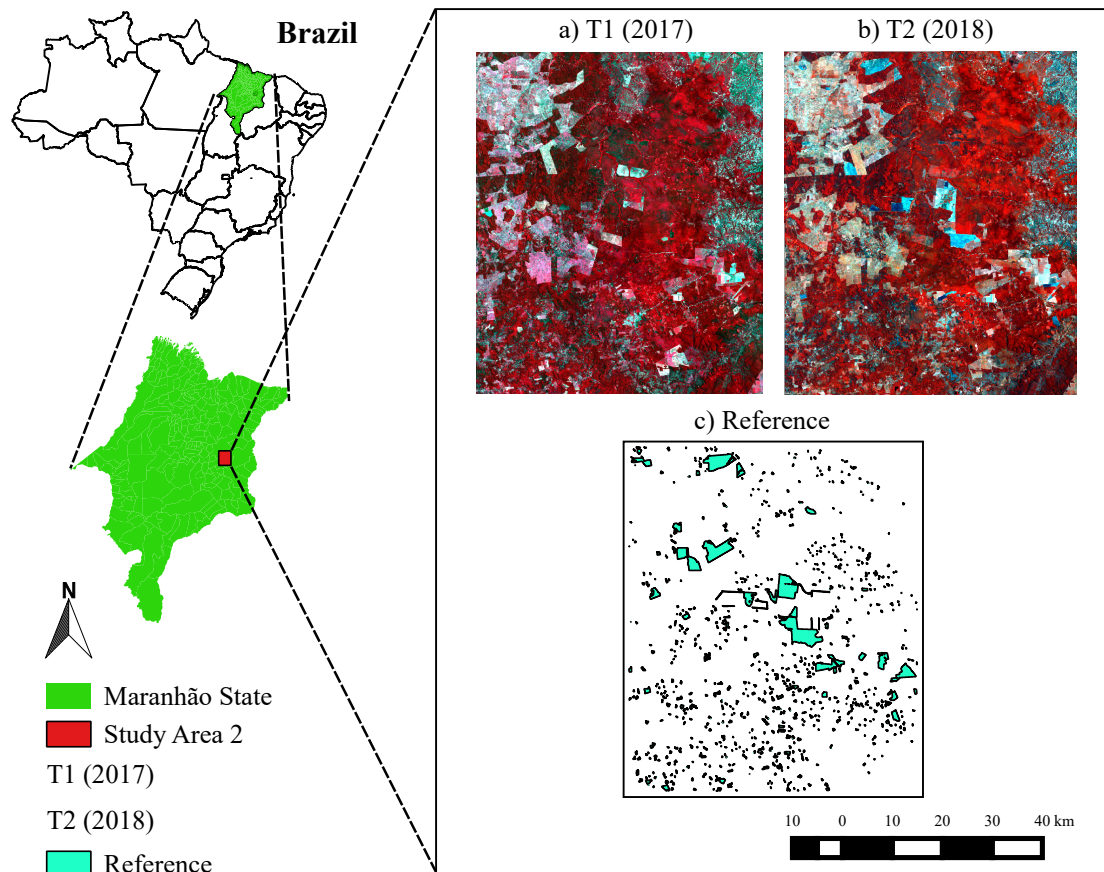


Figure 6. Region of the Brazilian Cerrado biome, located in Maranhão State, Brazil. *NIRGB* composition of the study area at two different dates T1 and T2 (a,b); (c) Reference of the deforestation process from 2017 to 2018.

- Some areas that suffered deforestation after the PRODES report were included in the reference. The added polygons were reviewed and approved by an expert photointerpreter. The final reference change map of the Cerrado is presented in Figure 6.
- An external buffer of two pixels around the samples of class “deforestation” was not considered in our evaluation to avoid the aforesaid inaccuracy problem along the borders.
- Areas lower than 1 ha (11 pixels) were not considered in the computation of the accuracy metrics because PRODES data does not consider deforested areas smaller than this value for the Cerrado biome.

2.5. Experimental Setup

For all the methods, two optical images acquired at different dates were used. Furthermore, the *Normalized Difference Vegetation Index (NDVI)* was computed (Equation (1)). This is an indicator of quantity and quality of vegetation, and it is measured from Landsat bands 5 and 4, which correspond to the near-infrared (*NIR*) and red (*Red*) ranges, respectively.

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

This index was appended to each image along the spectral dimension, so that the final images formed a tensor with depth equal to eight. Then, each individual spectral band was normalized to zero mean and unit variance.

The patch size was selected experimentally as 15×15 . Then, the input of EF and CSVM was a tensor of size of $15 \times 15 \times 16$, for SN a tensor of size of $15 \times 15 \times 8$ in each subnetwork and for SVM a vector of size of $15 \times 15 \times 16$. The procedure of the patch extraction was applied following the overlapping sliding windows with stride equal to three. The size of the patch and stride were selected empirically. In all methods, the input was an image patch, and the classification outcome was assigned to the patch central pixel.

Similar to [43], the images of Amazon and Cerrado databases were divided into 15 tiles, as shown in Figures 7 and 8, respectively. From each image, four tiles were selected for training, two tiles for validation, and nine tiles for testing.

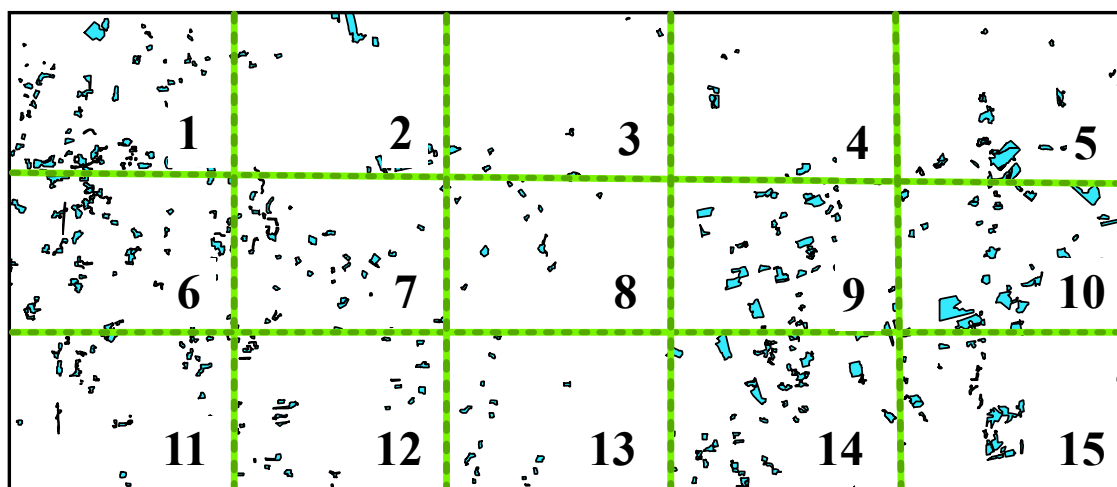


Figure 7. Distribution of the Amazon database. The region was divided into fifteen tiles. Four tiles were used for training (1, 7, 9, 13) and two for validation (5, 12). The remaining tiles were used for testing (2, 3, 4, 6, 8, 10, 11, 14, 15). The polygons indicate deforested areas.

As the number of available samples related with class “no-deforestation” was significantly higher (97% for Amazon biome and 95% for Cerrado biome) than class “deforestation” (3% for Amazon biome and 5% for Cerrado biome), a data augmentation on “deforestation” samples was adopted: 90° rotation, horizontal and vertical flip. To balance the number of samples per class, we further undersampled the class “deforestation”.

Tables 1 and 2 present the number of available patch pairs for training, validation, and test for the Amazon and Cerrado databases. Tables also present the number of patches obtained after applying the balancing procedure for both classes, “deforestation” and “no-deforestation”.

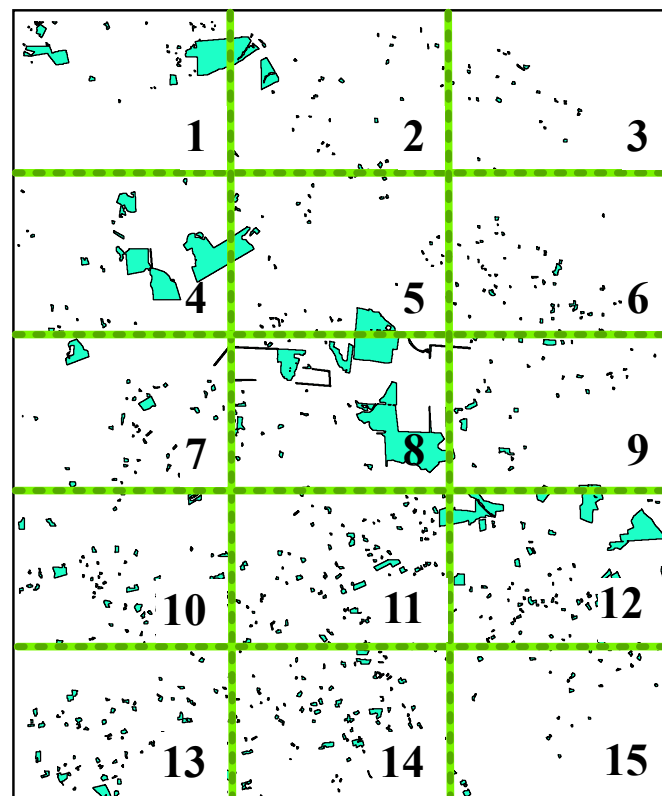
Table 1. Number of samples in the training, validation and test sets for Amazon database.

| Set | Tiles | Available Def. Samples | Available No-def. Samples | Balanced Samples (per Class) | Total Samples |
|------------|-------------------------------|------------------------|---------------------------|------------------------------|---------------|
| Training | 1, 7, 9, 13 | 2706 | 78,431 | 8118 | 16,236 |
| Validation | 5, 12 | 963 | 39,697 | 2889 | 5778 |
| Test | 2, 3, 4, 6, 8, 10, 11, 14, 15 | 40,392 | 1,675,608 | - | 1,716,000 |

The EF network architecture consisted of three convolutional layers (Conv) including the Rectified Linear Unit (ReLU), two Max-pooling layers (MaxPool), and two Fully Connected layers (FC), with a softmax layer at the end with two outputs, corresponding to “deforestation” and “no-deforestation” classes. The filter and output size of each layer are summarized in Table 3.

Table 2. Number of samples in the training, validation and test sets for Cerrado database.

| Set | Tiles | Available Def. Samples | Available No-def. Samples | Balanced Samples (per Class) | Total Samples |
|------------|------------------------------|------------------------|---------------------------|------------------------------|---------------|
| Training | 1, 5, 12, 13 | 4182 | 65,717 | 12,546 | 25,092 |
| Validation | 6, 10 | 663 | 34,658 | 1989 | 3978 |
| Test | 2, 3, 4, 7, 8, 9, 11, 14, 15 | 68,983 | 1,416,278 | - | 1,485,261 |

**Figure 8.** Distribution of the Cerrado database. The region was divided into fifteen tiles. Four tiles were used for training (1, 5, 12, 13) and two for validation (6, 10). The remaining tiles were used for testing (2, 3, 4, 7, 8, 9, 11, 14, 15). The polygons indicate deforested areas.**Table 3.** Architecture details of the EF model.

| Layer | Filter Size | Output Size | Parameters |
|-------------------|--------------|---------------------------|------------|
| Input | - | $15 \times 15 \times 16$ | - |
| Conv1 | 3×3 | $15 \times 15 \times 128$ | 18,560 |
| MaxPool1 | 2×2 | $7 \times 7 \times 128$ | - |
| Conv2 | 3×3 | $7 \times 7 \times 256$ | 295,168 |
| MaxPool2 | 2×2 | $3 \times 3 \times 256$ | - |
| Conv3 | 3×3 | $3 \times 3 \times 512$ | 1,180,160 |
| FC1 | - | 1×4608 | - |
| Dropout | - | 1×4608 | - |
| FC2 | - | 1×2 | 9218 |
| Total params | - | - | 1,503,106 |
| Treainable params | - | - | 1,503,106 |

Regarding the SN architecture, each subnetwork was also composed of three convolutional and two Max-pooling layers. The output of each subnetwork was fed to a FC layer and later concatenated in a single vector. In the end, a softmax layer generates the posterior probabilities for the classes

“deforestation” and “no-deforestation”. Table 4 shows the details of SN architecture with the filter and output size of each layer.

Table 4. Architecture details of the SN model.

| Layer | Filter Size | Output Size | Parameters |
|-------------------|--------------|---------------------------|------------|
| Input | - | $15 \times 15 \times 8$ | - |
| Conv1 | 3×3 | $15 \times 15 \times 128$ | 9344 |
| MaxPool1 | 2×2 | $7 \times 7 \times 128$ | - |
| Conv2 | 3×3 | $7 \times 7 \times 256$ | 295,168 |
| MaxPool2 | 2×2 | $3 \times 3 \times 256$ | - |
| Conv3 | 3×3 | $3 \times 3 \times 512$ | 1,180,160 |
| FC1 | - | 1×4608 | - |
| Concatenation | - | 1×9216 | - |
| Dropout | - | 1×9216 | - |
| FC2 | - | 1×2 | 18,434 |
| Total params | - | - | 1,503,106 |
| Treainable params | - | - | 1,503,106 |

For training the EF and SN models, we selected the following setup empirically: batch size equal to 32 with 100 number of epochs, early stopping after 10 epochs with no improvement (over the validation set) and a dropout rate of 0.2 in the final FC layer. Additionally, Adam optimizer was selected empirically with weight decay equal to 0.9 and learning rate equal to 10^{-3} . As loss function, we used the binary cross-entropy.

For the CSVM approach, the architecture comprised three convolutional layers, including ReLU, each one followed by a Max-pooling layer. The output size of each layer is shown in Table 5. In this method and for the baseline, the validation samples were added to the training set. For the computation of the weights of the SVM filters, the multicore Liblinear software package [50] was used. The parameter setup of the CSVM was: stride equal to one for the Conv and MaxPool layers, 12 SVMs used in each Conv layer. The training set was split in such a way that each SVM had the same number of samples for both classes. The size of the mini-patches used for learning the SVMs was equal to $3 \times 3 \times 16$ for the first convolutional layer and $3 \times 3 \times 12$ for the second and third layers. The estimation of the regularization parameter C for each SVM was performed using three-fold cross-validation restrained in the range $[10^{-1}, 10^3]$.

Table 5. Architecture details of the CSVM model.

| Layer | Filter Size | Output Size | Parameters |
|-------------------|--------------|--------------------------|------------|
| Input | - | $15 \times 15 \times 16$ | - |
| Conv1 | 3×3 | $13 \times 13 \times 12$ | 1740 |
| MaxPool1 | 1×1 | $11 \times 11 \times 12$ | - |
| Conv2 | 3×3 | $9 \times 9 \times 12$ | 1308 |
| MaxPool2 | 1×1 | $7 \times 7 \times 12$ | - |
| Conv3 | 3×3 | $5 \times 5 \times 12$ | 1308 |
| MaxPool3 | 1×1 | $3 \times 3 \times 15$ | - |
| Total params | - | - | 4356 |
| Treainable params | - | - | 4356 |

The buffer of both references was obtained applying the morphological dilation, using as structuring elements a disk of radius 2. This operation expanded the boundaries of the deforested polygons. Then, a difference between the dilated and original images was performed, resulting in the outer edge, and the patches with the central pixel in these regions were not considered for training, validation or test.

2.6. Influence of the Number of Training Samples

To evaluate the influence of the number of training samples, four scenarios were considered. Specifically collecting samples from the training set of a one, two, three and four tiles, denoted as N_i , where i corresponds to the number of tiles used in each scenario. For EF and SN methods, the validation set (val) was used to stop training once the loss increased in 10 consecutive epochs (early stopping). As mentioned before, for CSVM and SVM the samples in this set were added to the training set (tr). The number of training samples in each scenario for the Amazon and Cerrado databases is presented in Tables 6 and 7, respectively.

Table 6. Training tiles used for the Amazon database.

| Training Set | Tiles | Available Def. Samples | Available No-def. Samples | Balanced Samples (per Class) | Total Samples (tr + val) |
|--------------|-------------|------------------------|---------------------------|------------------------------|--------------------------|
| 1 Tile | 13 | 239 | 20,306 | 717 | 1434 + 5778 |
| 2 Tiles | 1, 13 | 709 | 40,515 | 2127 | 4254 + 5778 |
| 3 Tiles | 1, 7, 13 | 1807 | 59,102 | 5421 | 10,842 + 5778 |
| 4 Tiles | 1, 7, 9, 13 | 2706 | 78,431 | 8118 | 16,236 + 5778 |

Table 7. Training tiles used for the Cerrado database.

| Training Set | Tiles | Available Def. Samples | Available No-def. Samples | Balanced Samples (per Class) | Total Samples (tr + val) |
|--------------|--------------|------------------------|---------------------------|------------------------------|--------------------------|
| 1 Tile | 5 | 671 | 17,370 | 2013 | 4026 + 3,978 |
| 2 Tiles | 5, 13 | 1240 | 33,760 | 3720 | 7440 + 3,978 |
| 3 Tiles | 1, 5, 13 | 2287 | 50,273 | 6861 | 13,722 + 3,978 |
| 4 Tiles | 1, 5, 12, 13 | 4182 | 65,717 | 12,546 | 25,092 + 3978 |

2.7. Accuracy Assessment

The performance of the evaluated methods was expressed in terms of *Overall Accuracy (OA)*, *F1-Score*, and *Alarm Area (AA)*.

- *Overall Accuracy (OA)*: is a global metric that indicates the percentage of samples correctly classified in relation to the total samples. It is defined by:

$$OA = \frac{tp + tn}{P + N} \times 100 \quad (2)$$

where *true positives (tp)* is the number of samples correctly assigned to the class “deforestation”, *false positives (fp)* refer to the number of samples erroneously assigned to the class “deforestation”. Analogously, *true negatives (tn)* and *false negatives (fn)* correspond to the number of samples correctly and incorrectly assigned to the class “no-deforestation”, respectively. P and N denote the total number of positive and negative samples in the test set.

- *Precision*, also known as *Correctness*, represents the proportion of samples assigned by the classifier to the class “deforestation”, which truly belongs to that class, formally

$$Precision = \frac{tp}{tp + fp} \quad (3)$$

- *Recall*, also known as *Completeness*, is the proportion of all “non deforestation” samples recognized by the classifier as such, i.e.,

$$Recall = \frac{tp}{tp + fn} \quad (4)$$

- *F1-score*: is given by the harmonic mean of *Precision* and *Recall* and it also varies in a range of 0 to 1. This metric is defined by:

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100 \quad (5)$$

- *Alarm Area*: this metric is the portion of the monitored area classified as “deforestation”. We defined this metric by the rate of *tp* and *fp* between the total *P* and *N* samples in the test set.

$$AA = \frac{tp + fp}{P + N} \times 100 \quad (6)$$

This metric is important in an operational scenario where an automatic system highlights areas suspected of deforestation (alarm), which will be subsequently evaluated visually by a human analyst to eliminate false positives. The lower the *Alarm Area*, the lower is the human effort.

3. Results and Discussion

In this section, we present and discuss the results obtained by the methods described in Section 2 for the Amazon and Cerrado databases. Firstly, we report the average of *Overall Accuracy (OA)* and *F1-score* computed over ten runs, each run with a different choice of training samples for the class “no deforestation”. Next, we present the probability maps generated in each experiment, and finally, we analyze how semi-automatic approaches could be designed based on these methods to reduce human intervention with minimal accuracy loss.

3.1. Amazon Biome

Figure 9 summarizes the results of the experiments on the Amazon biome in respect of *F1-score* for the class “deforestation”. This figure shows that SN and EF achieved the best performance in most results. According to expectations, the performance of all methods improved as more patches were used for training. However, CSVM was only able to reach the baseline performance when four tiles were used for training. It attained low scores in comparison with the other methods. With a single training tile, the CSVM performance was similar for two and three layers; with two and four tiles for training, the best performance was obtained with two layers; using three tiles for training, the performance decreased as more layers were added.

The SN was the best performing method, followed by SVM and then EF. This was not surprising, given the well-known generalization capacity of SVM in the face of the scarce training data. Contrarily, SN and EF consistently outperformed SVM in about 10% and 13% when data from two, three, and four tiles were used for training.

The results in terms of *Overall Accuracy (OA)* are presented in Figure 10. Similar to the *F1-score*, OA improved as more training samples were available. In all cases, OA scores over 90% were obtained. Again, CSVM presented lower scores in comparison to other methods, and its performance was similar to *F1-score* in each scenario. The high OA scores can be understood because about 97% of the test samples were from class “no-deforestation”.

Figures 11 and 12 show the *NIR-G-B* composition (Near Infrared, Green, and Blue bands) at both dates (T1 and T2), the reference, and the probability maps for tiles 2 and 14, respectively. These tiles are part of the test set. Columns correspond to methods SVM, EF, SN, and CSVM (after layer 2), and rows correspond to using one, two, three, and four tiles for training. Blue color represents the lowest probability of deforestation, while the red color represents the highest probability.

As in the *F1-score* and OA plots, the probability maps improved, and the salt-and-pepper effect reduced when the number of training tiles augmented. In the first scenario, when a single tile was used for training, SVM delivered many false positives, followed by CSVM, causing a more noticeable salt-and-pepper effect.

SVM was least confident among the tested methods in its results. It produced comparatively many intermediate probability values, whereas its counterparts generated probabilities more concentrated close to 0 and 1. All methods presented intermediate probabilistic values mainly around polygon borders. Inaccuracies in the reference of deforested polygons might have contributed to this behavior.

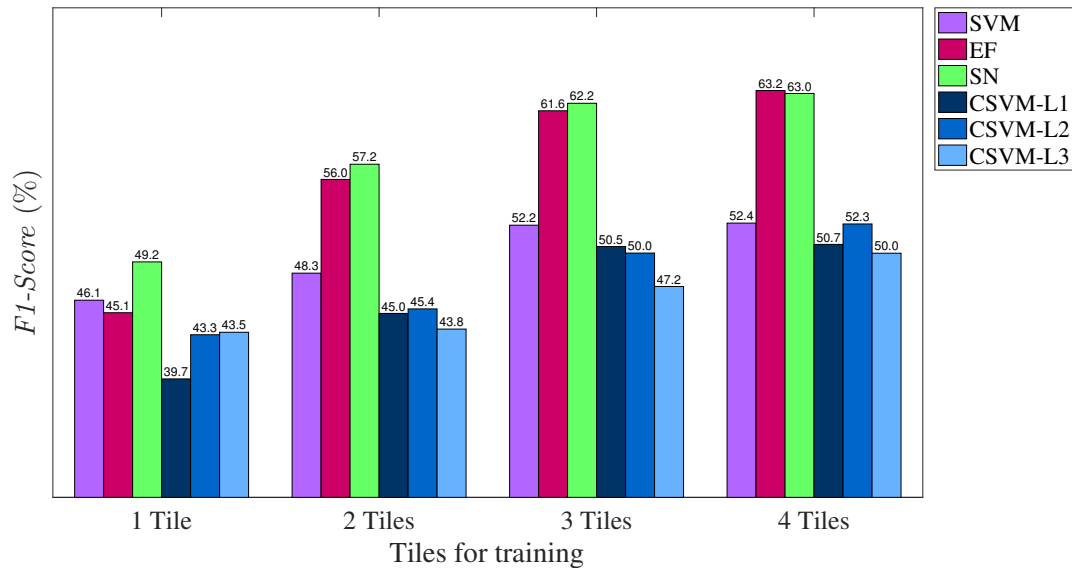


Figure 9. F1-score of Amazon Biome obtained from SVM, EF, SN, and CSVM (after each layer) using one, two, three, and four tiles for training.

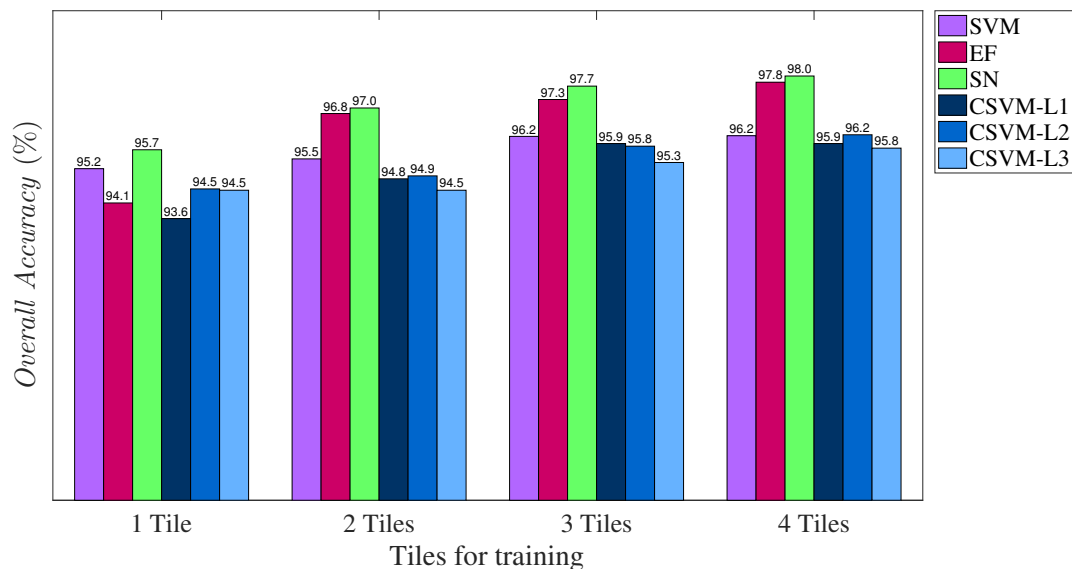


Figure 10. Overall Accuracy of Amazon Biome obtained from SVM, EF, SN, and CSVM (after each layer) using one, two, three, and four tiles for training.

3.2. Alarm Area vs. Recall for Amazon Biome

Next, we evaluate the methods as part of an alarm scheme. In this scheme, the underlying classifier indicates areas where deforestation is likely to have occurred. A photointerpreter then visually analyzes the image, or an inspector could be sent to the indicated areas to check what was real deforestation and what was just a false alarm. The main benefit of this scheme is to restrict the human effort to just a portion of the area being monitored.

On the other hand, in this scheme, parts of the deforested areas can be undetected by the classifier and go unnoticed. Two metrics are critical in this analysis: first, the proportion of monitored area

flagged as potentially deforested and, second, the proportion of total deforestation concentrated in the areas indicated by the classifier. The first metric is the *Alarm Area* defined in Equation (6), whereas the second metric is the *Recall* defined in Equation (4).

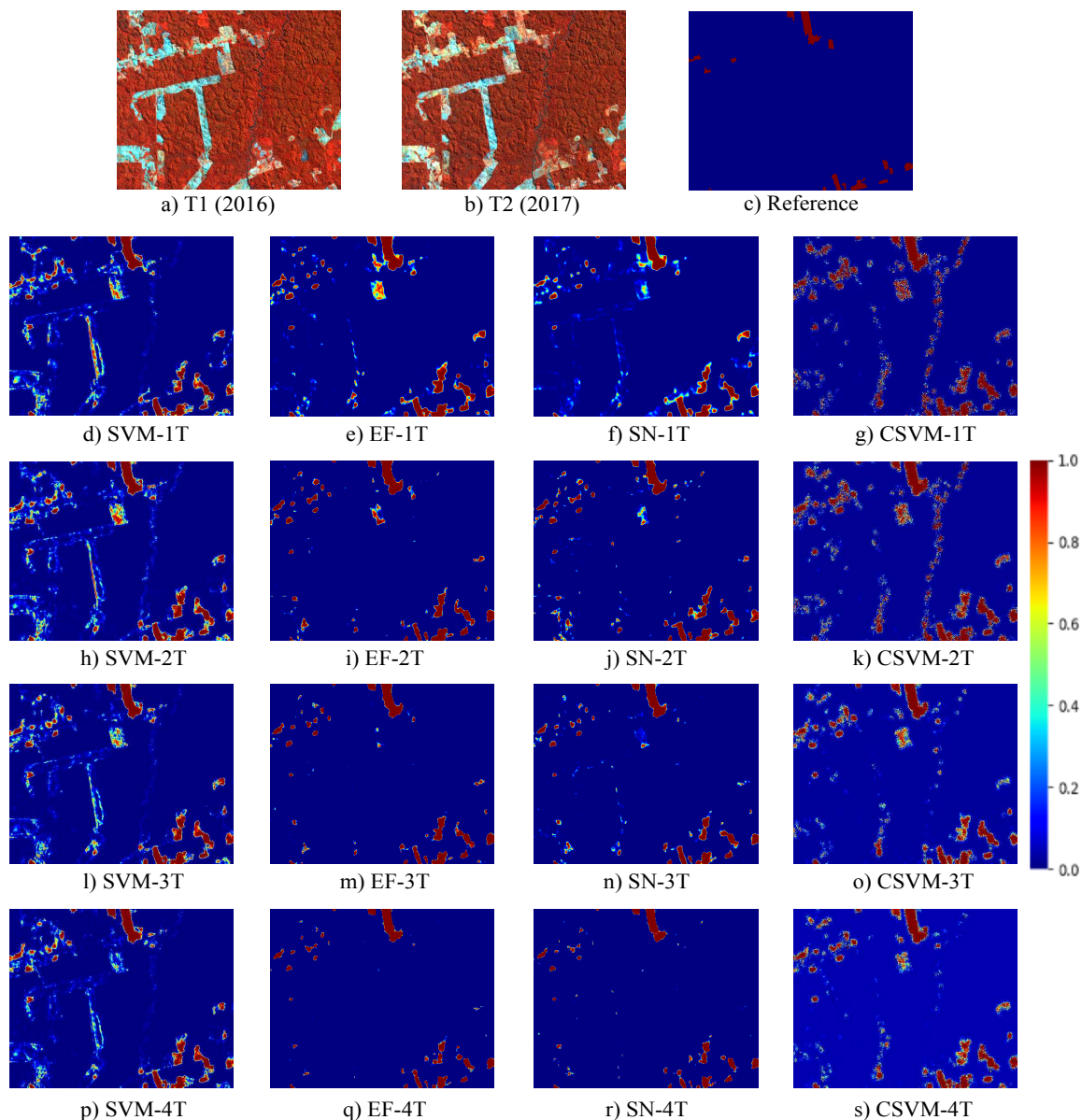


Figure 11. Predicted maps for tile 2 computed by SVM, EF, SN, and CSVM using one, two, three, and four tiles for training.

Both metrics will depend on a threshold for the deforestation probability assigned by the classifier above which a site should deserve attention. The higher this threshold, the smaller the *Alarm Area* and the smaller the *Recall*. The threshold value expresses a tradeoff between accuracy and human effort and will be determined by the operational demands at each time and each region. Therefore, the following analysis focuses on the behavior of these two metrics as the deforestation probability threshold varies.

Specifically, we present the curves *Recall* versus *Alarm Area* for each method. Each point in the curve corresponds to a threshold imposed on the deforestation probability produced by each tested method. A small area to be checked out at a high *Recall*, is the desired profile.

With one tile for training, all methods achieved *Recall* values of about 90% when looking at less than 10% of the whole imaged area. It means that 90% of the correctly identified deforestation is

contained in 10% of the image. Hence, instead of looking at the entire image, the analyst would focus on 10% of it, reducing human work by 90%. As expected, as *Recall* increased, the area to be observed also increased, but, in this particular case, CSVM (with three layers) presented the best results (see Figure 13a). For *Recall* beyond 96%, the threshold values were very close to zero, most pixels tended to be classified as deforestation, and the area to be observed went up to 100%, as can be observed in Figure 13b.

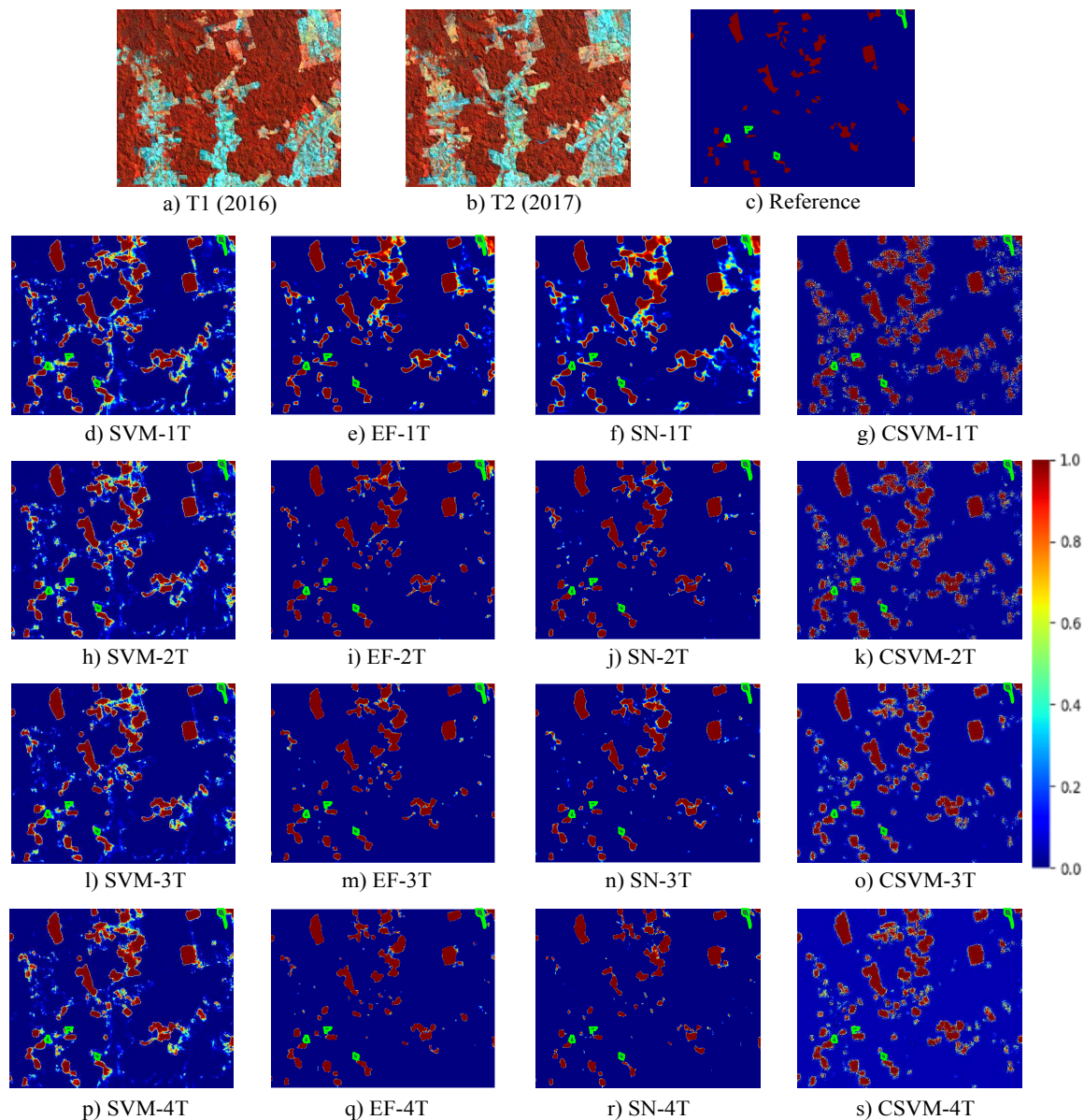


Figure 12. Predicted maps for tile 14 computed by SVM, EF, SN, and CSVM using one, two, three, and four tiles for training.

Using two, three, and four tiles for training, all methods presented a similar profile until 96% *Recall* but, beyond this value, SVM presented the best performance. It managed to classify more deforested samples correctly, with a minimum increase in the area to be observed. Analogous to the results for one training tile, when threshold values were set very close to zero, all samples tended to be classified as deforestation, and the *Recall* approached quickly to 100%, as well as the area to be observed.

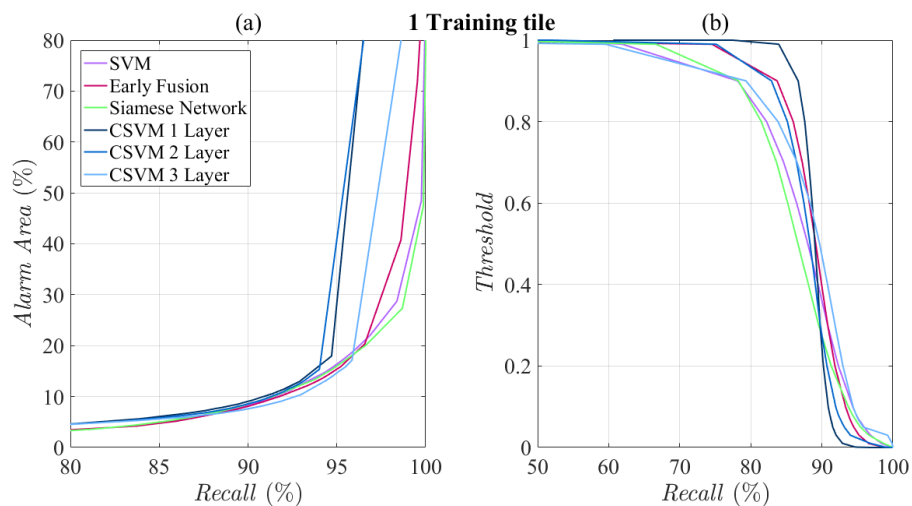


Figure 13. (a) Recall vs. Alarm Area (deforested area) for Amazon biome using one training tile. (b) The recall values are obtained by threshold variation.

A somewhat surprising conclusion from Figures 13–16 is that the CSVM performed close to the other methods. Contrarily, SVM performed significantly worse than EF and SN in the analysis reported in the preceding section. Notice that, in that case, we implicitly set the probability threshold to 50% to discriminate the classes. In the present analysis in which the threshold varies, there was no significant superiority of the other methods over the SVM. The experiments, therefore, indicated that SVM might be an attractive option for an alarm system, given its low demand for training samples when compared to DL-based methods.

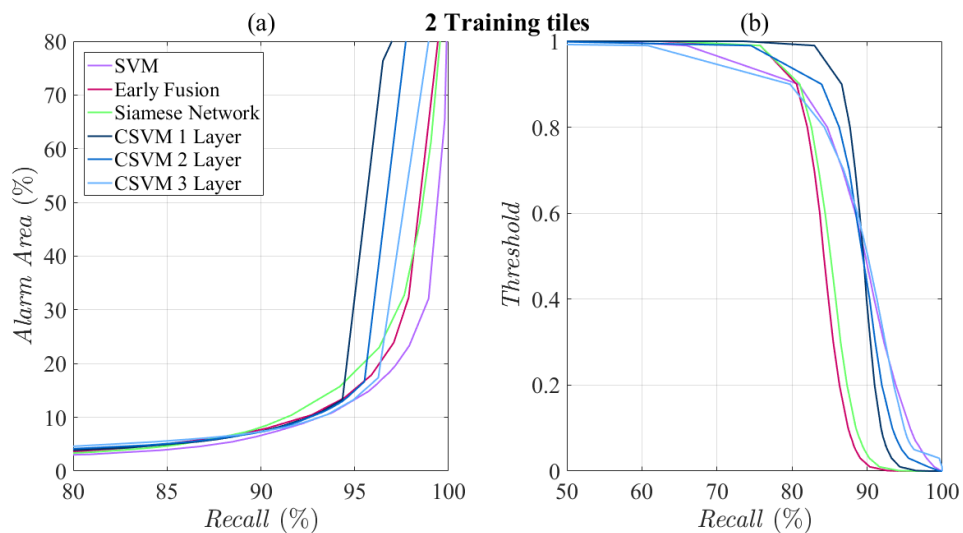


Figure 14. (a) Recall vs. Alarm Area (deforested area) for Amazon biome using two training tiles. (b) The recall values are obtained by threshold variation.

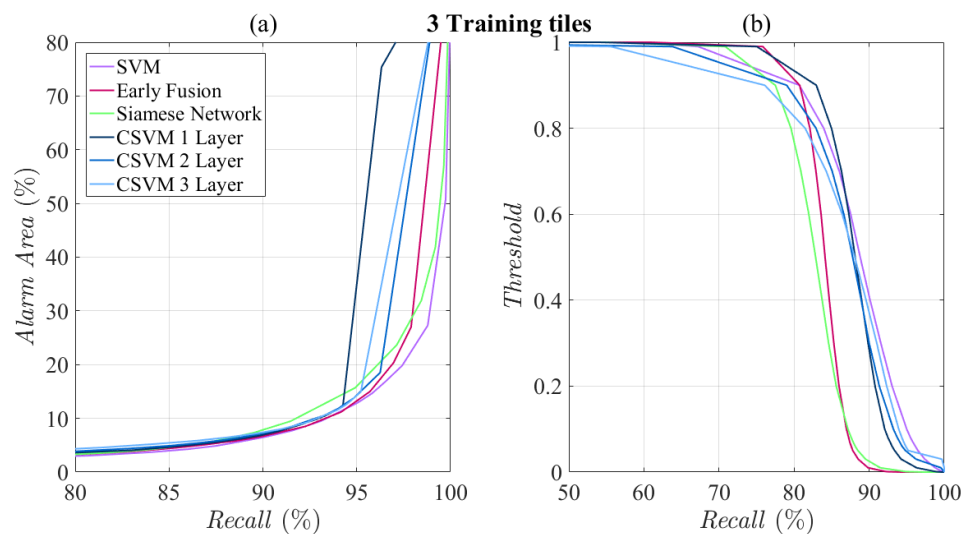


Figure 15. (a) Recall vs. Alarm Area (deforested area) for Amazon biome using three training tiles. (b) The recall values are obtained by threshold variation.

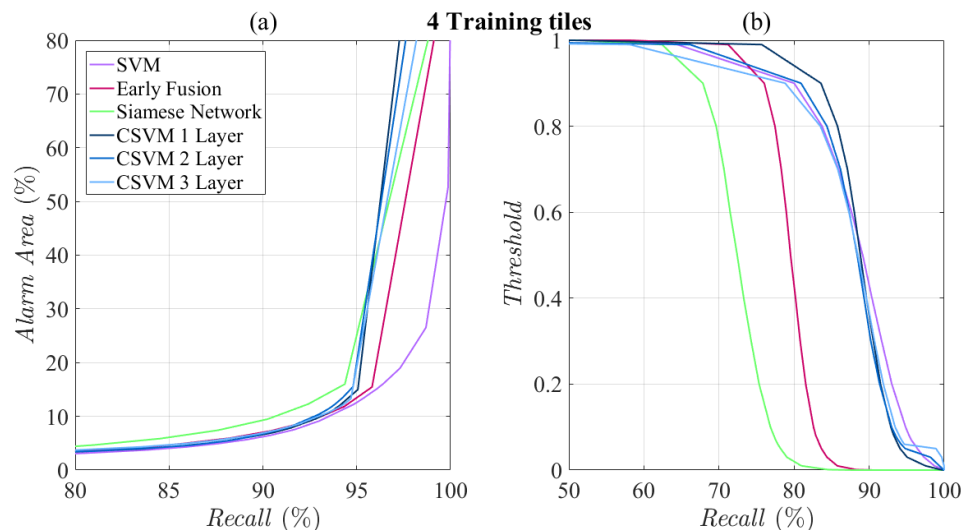


Figure 16. (a) Recall vs. Alarm Area (deforested area) for Amazon biome using four training tiles. (b) The recall values are obtained by threshold variation.

3.3. Cerrado Biome

The results for Cerrado in terms of *F1-score* and *OA* are summarized in Figures 17 and 18, respectively. Similar to the Amazon database, EF and SN presented the best performance in all experiments. Using a single tile for training, EF and SN outperformed SVM in 2% and 3% respectively. The best performance achieved by CSVM was 51%, which was obtained after the first layer. However, it did not reach the baseline. Using two tiles for training, EF and SN outperformed SVM with a difference of about 2%. In this case, SVM outperformed CSVM by 9%. Using three tiles for training, EF and SN outperformed SVM in 3% and 2%, respectively, and CSVM in the second layer came very close to SVM. Using four training tiles, the DL-based methods were better than SVM. EF and SN and CSVM (one layer) overcome SVM in 2%, 3%, and 1%, respectively.

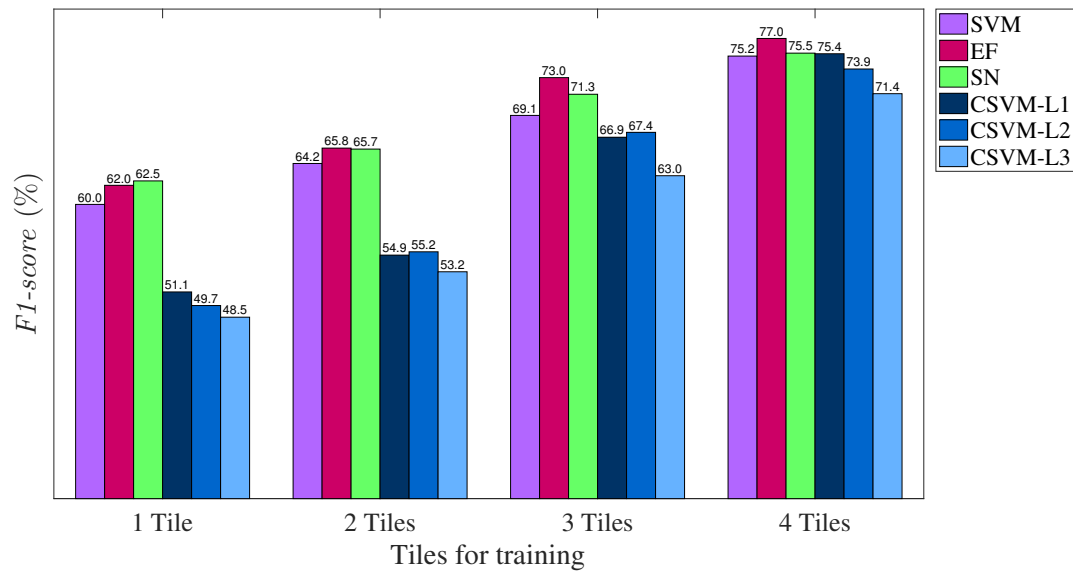


Figure 17. *F1-score* of Cerrado Biome obtained from SVM, EF, SN, and CSVM using one, two, three, and four training tiles.

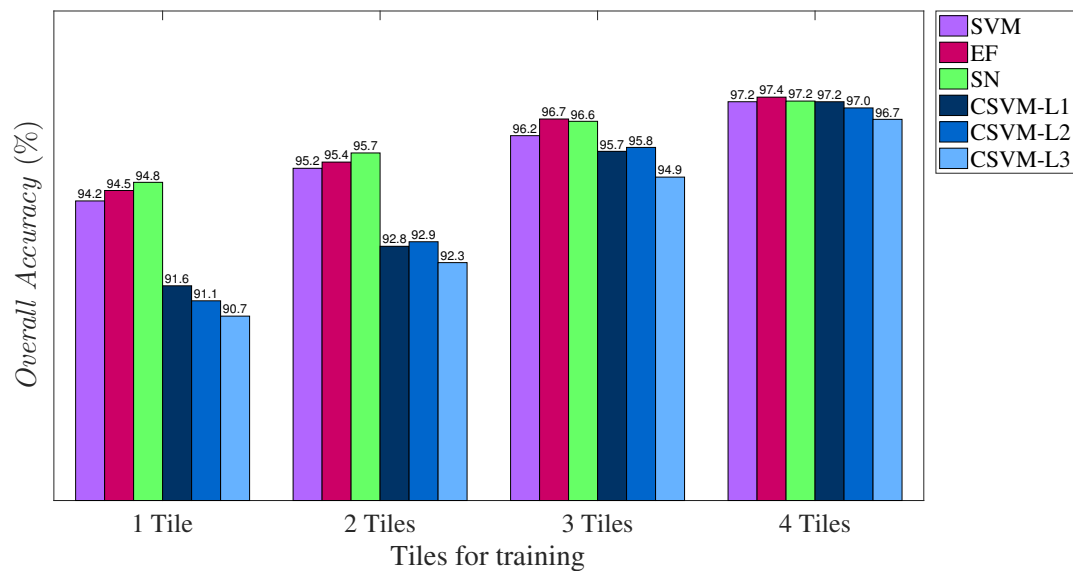


Figure 18. Overall Accuracy of Cerrado Biome obtained from SVM, EF, SN, and CSVM (after each layer) using one, two, three, and four tiles for training.

In terms of OA, the results presented a similar trend observed on *F1-score*. Scores above 90% were achieved in all scenarios. However, EF and SN obtained the best performance in all experiments. Analogous to the experiments on the Amazon database, CSVM presented lower scores in comparison with other methods. Only in the last case, using four training tiles, CSVM matched SVM at 97%. As in the experiments in the Amazon dataset, the high OA values were because the vast majority of samples belonging to the class “no-deforestation” were correctly classified.

Figures 19 and 20 show the *NIR-G-B* composition (Near Infrared, Green, and Blue bands) at both dates (T1 and T2), the reference, and the probability maps of tile 2 and 8, respectively. Again, columns correspond to methods SVM, EF, SN, and CSVM (after layer 1), and rows correspond to the results one, two, three, and four training tiles. Blue represents the lowest probability of deforestation, while Red represents the highest probability.

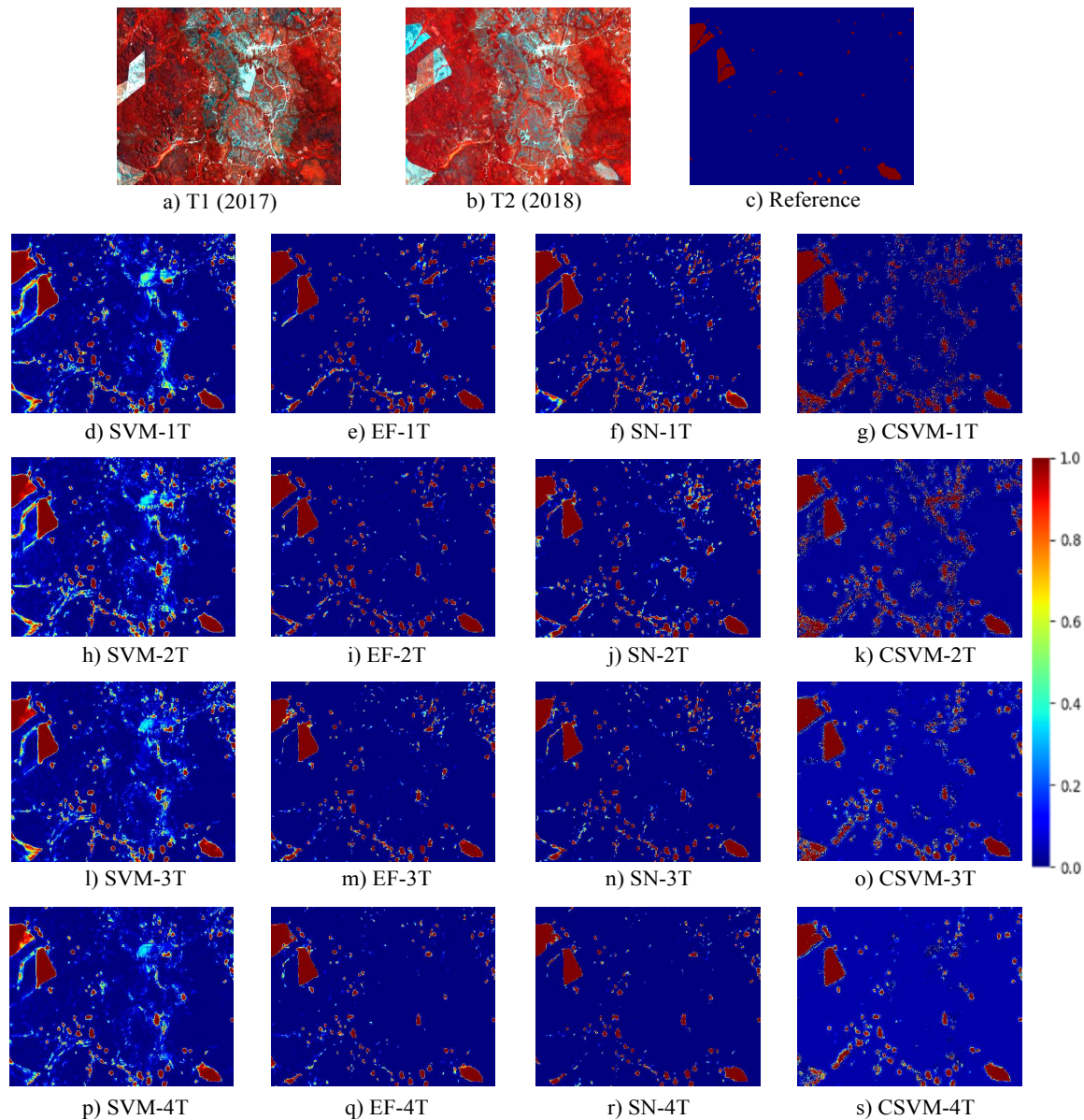


Figure 19. Predicted maps of tile 2 produced by SVM, EF, SN, and CSVM using one, two, three, and four tiles for training.

Like the results recorded on the Amazon database, the probability maps improved, and the salt-and-pepper effect reduced as the number of training tiles increased. If we observe the first scenario, where a single tile was used for training, all maps present a large number of false positives and a notable salt-and-pepper effect. Likewise, EF, SN, and CSVM are more confident, assigning values close to one for pixels of class “deforestation”, and values close to zero to pixels of “no-deforestation” class. Contrarily, the probability maps delivered by SVM contain comparatively many pixels with probability values in the intermediate range.

As observed in the previous experiment series, the probability maps show that all methods were less confident, i.e., present probability values around 50%, close to the borders of the reference polygons. As mentioned before, this is possibly related to inaccuracies in the delimitation of deforestation polygons in the reference.

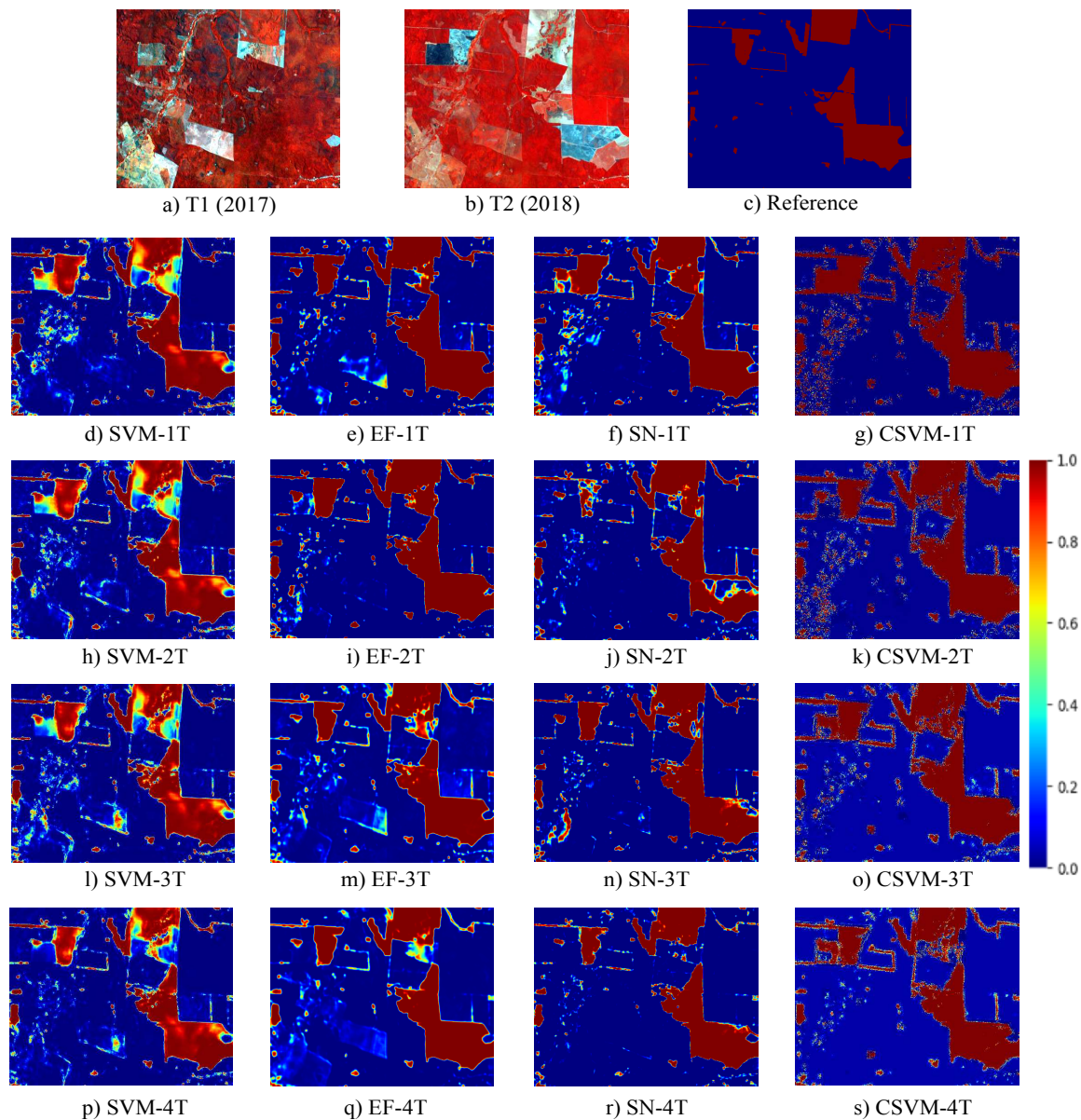


Figure 20. Predicted maps tile 8 produced by SVM, EF, SN, and CSVM using one, two, three, and four tiles for training.

3.4. Alarm Area vs. Recall for Cerrado Biome

The analysis under the perspective of an alarm system is presented in Figures 21–24 for one, two, three, and four training tiles, respectively. For this database, in the four scenarios, the best performance was obtained by EF. Although with a single training tile, the performance was similar for all methods, EF was slightly superior. Using two tiles for training, EF, was also the best performing method, followed by CSVM and SVM, which presented very similar results. Finally, using three and four tiles for training, EF and CSVM achieved better results: they correctly classified more samples of class “deforested” and the area to be observed is lower. According to the graphs at 95% of *Recall*, the area to be observed is reduced to 10% of the entire image (see Figures 21a, 22a, 23a and 24a). In the same case of the Amazon database, for threshold values close to zero, all samples are classified as deforestation class, the value of *Recall* is about 100% and the area to be observed is the entire image, as can be seen in Figures 21b, 22b, 23b and 24b.

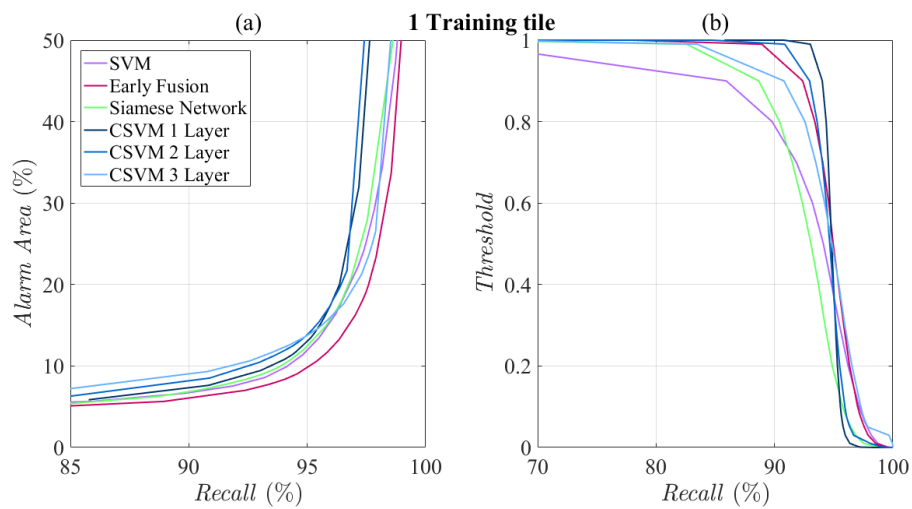


Figure 21. (a) Recall vs. Alarm Area (deforested area) for Cerrado biome using one training tile. (b) The recall values are obtained by threshold variation.

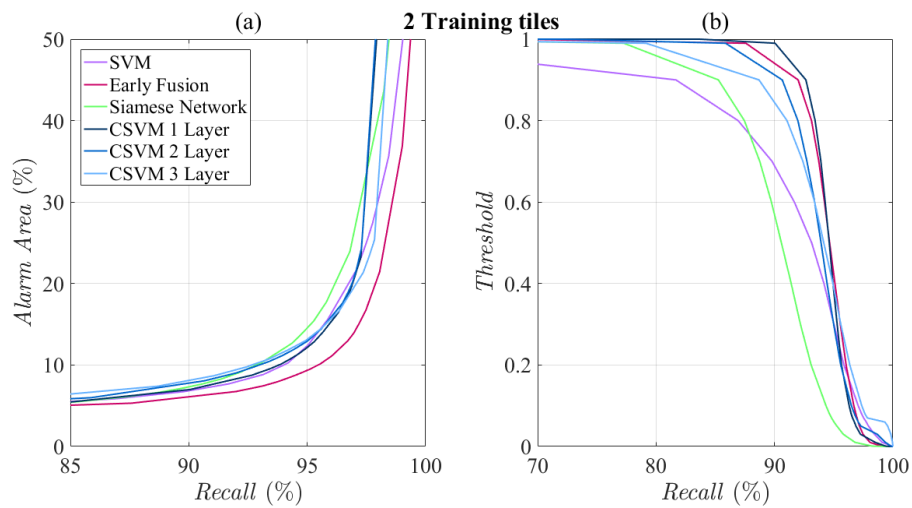


Figure 22. (a) Recall vs. Alarm Area (deforested area) for Cerrado biome using two training tiles. (b) The recall values are obtained by threshold variation.

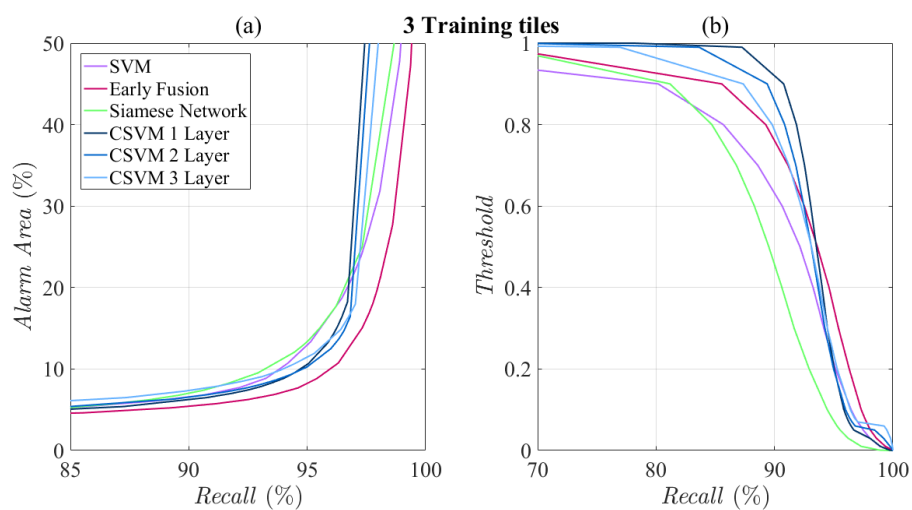


Figure 23. (a) Recall vs. Alarm Area (deforested area) for Cerrado biome using three training tiles. (b) The recall values are obtained by threshold variation.

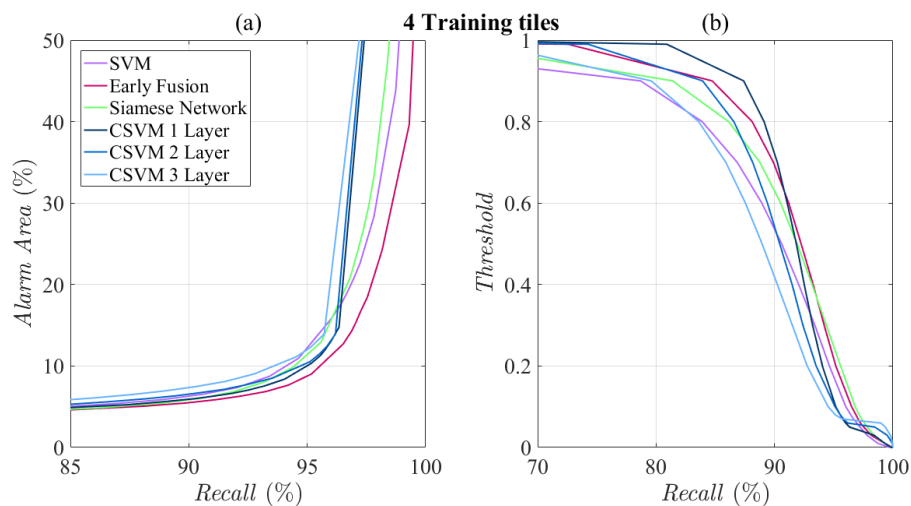


Figure 24. (a) Recall vs. Alarm Area (deforested area) for Cerrado biome using four training tiles. (b) The recall values are obtained by threshold variation.

Compared with the results of the analysis conducted on the Amazon biome, the superiority of EF over its competitors was more pronounced here, especially for *Recall* values starting at 95%, even when only one training tile was used.

4. Conclusions

This work reported an evaluation of three state-of-the-art deep learning techniques for deforestation detection: Early Fusion (EF), Siamese Network (SN), and Convolutional SVM (CSVM). Additionally, the performance of these methods was compared against a baseline based on probabilistic Support Vector Machine (SVM), which is one of the most popular machine learning techniques for change detection.

Experiments were carried out using two areas of the Brazilian biomes. The first one corresponds to a region of the Amazon biome, and the second corresponds to the Cerrado biome. The references used in this work were collected from the PRODES Project, which was developed by the National Institute for Space Research (INPE). The methodology employed to accomplish this task involves significant human intervention. This work has a great potential at reducing human intervention and assessing state-of-the-art methods towards more automatic deforestation detection. With the improvements resulting from the proposed techniques, the mapping can be performed in less time, with lower costs and with a lower degree of subjectivity.

The experimental analysis relied on two LANDSAT 8/OLI optical images acquired at dates about one year apart from each other. Four different scenarios were considered, using one, two, three, and four tiles training. As expected, the performance of all methods increased with the number of training samples. This trend was chiefly remarkable for EF and SN.

EF and SN presented the best performance in most experiments. In a few cases, CSVM outperformed SVM. The accuracy obtained by EF and SN in experiments were up to 95% in terms of *Overall Accuracy (OA)* and up to 63% in terms of *F1-score* for Amazon, and up to 97% in terms of *OA* and 78% in terms of *F1-score* for Cerrado, showing that the results for the Cerrado database achieved a higher percentages than Amazon database. The reason lies in the pattern of deforestation in the Cerrado biome. It is comparatively more intense; the vegetation is completely removed, and most of the soil is exposed, unlike the Amazon, where it is common to have vegetation remains in the deforestation process, which hinders detection.

Besides, the probability maps indicated that EF, SN, and CSVM were more confident in their outcomes. Most posterior probabilities delivered by these methods were concentrated close to one and

zero, for deforestation and no-deforestation, respectively, whereas the posteriors computed by SVM took intermediate values over comparatively many areas.

The main motivation for including CSVM in this study was the good performance reported in a recent paper under a small training sample size. The experimental analysis did not confirm this expectation. Indeed, in our experiments, CSVM was consistently outperformed by EF and SN, and in few cases, also by SVM.

Regarding CSVM, some additional experiments for the final classification layer were performed. The first one was the usage of the flattening feature maps obtained after each convolutional layer to train the binary SVM, instead of pooling them over four quadrants and calculate the means. The second experiment involved the selection of the final classifier. We tested a Softmax layer classifier and SVM with an RBF (Radial Basis Function) and a linear kernel. However, the best results were obtained using a linear kernel.

It is worth to mention that despite CSVM did not overcome the baseline in most cases, it presents an advantage concerning EF and SN, it is a CPU-based method, then it does not require GPU to carry out the experiments. GPU is much more costly, and it relies on powerful supplementary equipment to support it.

An additional evaluation was also performed to verify how the methods can reduce the time invested in the visual inspection of deforested areas. The metric defined as *Alarm Area* (AA) was computed to evaluate how the methods can reduce the human analyst effort for visual classification, by restricting the whole image to just a portion of the total area being monitored. According to the experiments, for the Amazon biome, it was estimated that it would be possible to reduce the human work by 90%, with the guarantee that 90% of the deforestation occurrences are present in 10% of the whole imaged area. For the Cerrado biome, 95% of the deforestation occurrences would be present in a similar portion of the image.

Although the evaluated methods were tested on deforestation detection, they can be easily adapted to other change detection applications. In the present study, these methods proved to be promising directions in the research to monitor and control environmental issues that are of paramount importance today.

Further studies should test other combinations of hyperparameters of the assessed methods with a focus on decreasing the number of false deforestation samples, as well as to evaluate other deep architectures, such as Fully Convolutional Networks (FCN) and Recurrent Neural Networks (RNN). Another investigation direction relates to the usage of freely available data from other sensors. An attractive alternative is the Sentinel-2 data that provides better temporal and spatial accuracy than LANDSAT-8. Furthermore, the management of the Synthetic Aperture Radar (SAR), would allow monitoring deforestation in a way that is almost independent of weather conditions. Indeed, the Brazilian biomes present a cloud coverage for nearly the whole year, which prevents the use of optical imagery. Given these circumstances, SAR images are a promising option.

A critical issue is still the number of training samples required by deep learning-based methods to achieve their full potential. Techniques based on domain adaptation seem another promising research direction to mitigate this hindrance.

Author Contributions: Experiments and writing, original draft preparation, M.O.A.; overall study design, R.Q.F.; co-supervision, writing, review and editing, P.N.H., C.A.D.A. and A.R.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Acknowledgments: The authors are grateful to Professor Farid Melgani and his research group for providing the original CSVM code.

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

References

- De Sy, V.; Herold, M.; Achard, F.; Beuchle, R.; Clevers, J.; Lindquist, E.; Verchot, L. Land use patterns and related carbon losses following deforestation in South America. *Environ. Res. Lett.* **2015**, *10*, 124004.
- MacDicken, K.; Jonsson, Ö.; Piña, L.; Maulo, S.; Contessa, V.; Adikari, Y.; Garzuglia, M.; Lindquist, E.; Reams, G.; D'Annunzio, R. *Global Forest Resources Assessment 2015: How Are the World's Forests Changing?* Food and Agricultural Organization of the United Nations (FAO): Rome, Italy, 2015.
- Amin, A.; Choumert-Nkolo, J.; Combes, J.L.; Motel, P.C.; Kéré, E.N.; Ongono-Olinga, J.G.; Schwartz, S. Neighborhood effects in the Brazilian Amazônia: Protected areas and deforestation. *J. Environ. Econ. Manag.* **2019**, *93*, 272–288. [[CrossRef](#)]
- Aide, T.M.; Clark, M.L.; Grau, H.R.; López-Carr, D.; Levy, M.A.; Redo, D.; Bonilla-Moheno, M.; Riner, G.; Andrade-Núñez, M.J.; Muñoz, M. Deforestation and Reforestation of Latin America and the Caribbean (2001–2010). *Biotropica* **2013**, *45*, 262–271. [[CrossRef](#)]
- Brazilian Institute of Geography and Statistics (IBGE). Mapa de Biomassas e de Vegetação. 2004. Available online: <https://ww2.ibge.gov.br/home/presidencia/noticias/21052004biomasshtml.shtm> (accessed on 20 November 2019).
- Almeida, C.A.A.D.; Coutinho, A.C.; Esquerdo, J.A.C.A.D.M.; Adami, M.; Venturieri, A.; Diniz, C.G.; Dessay, N.; Durieux, L.; Gomes, A.R. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. *Acta Amaz.* **2016**, *46*, 291–302. [[CrossRef](#)]
- Sano, E.E.; Rosa, R.; Brito, J.L.A.S.S.; Ferreira, L.G.A. Mapeamento semidetalhado do uso da terra do Bioma Cerrado. *Pesqui. Agropecuaria Bras.* **2008**, *43*, 153–156. [[CrossRef](#)]
- Soterroni, A.C.; Ramos, F.M.; Mosnier, A.; Fargione, J.; Andrade, P.R.; Baumgarten, L.; Pirker, J.; Obersteiner, M.; Kraxner, F.; Câmara, G.; et al. Expanding the Soy Moratorium to Brazil's Cerrado. *Sci. Adv.* **2019**, *5*, doi:10.1126/sciadv.aav7336. [[CrossRef](#)]
- Bueno, I.T.; Acerbi Júnior, F.W.; Silveira, E.M.; Mello, J.M.; Carvalho, L.M.; Gomide, L.R.; Withey, K.; Scolforo, J.R.S. Object-Based Change Detection in the Cerrado Biome Using Landsat Time Series. *Remote Sens.* **2019**, *11*, 570. [[CrossRef](#)]
- Goodman, R.; Aramburu, M.; Gopalakrishna, T.; Putz, F.; Gutiérrez, N.; Alvarez, J.; Aguilar-Amuchastegui, N.; Ellis, P. Carbon emissions and potential emissions reductions from low-intensity selective logging in southwestern Amazonia. *For. Ecol. Manag.* **2019**, *439*, 18–27. [[CrossRef](#)]
- Malingreau, J.; Eva, H.; De Miranda, E. Brazilian Amazon: a significant five year drop in deforestation rates but figures are on the rise again. *Ambio* **2012**, *41*, 309–314. [[CrossRef](#)]
- Barreto, P.; Souza, C., Jr.; Nogueira, R.; Anderson, A.; Salomão, R. *Human Pressure on the Brazilian Amazon Forests*; World Resources Institute: Washington, DC, USA, 2006.
- National Institute for Space Research (INPE). Monitoring of the Brazilian Amazonian Forest by Satellite; 1988. Available online: <http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes> (accessed on 5 November 2019).
- World Wildlife Fund (WWF). Amazon Deforestation. 1988. Available online: http://wwf.panda.org/our_work/forests/deforestation_fronts/deforestation_in_the_amazon/ (accessed on 7 December 2019).
- Ministry of the Environment (MMA); Brazilian Institute of Environment and Renewable Natural Resources (IBAMA). Monitoramento do desmatamento nos biomas brasileiros por satélite. 2011. Available online: http://www.mma.gov.br/estruturas/sbf_chm_rbbio/_arquivos/relatoriofinal_cerrado_2010_final_72_1.pdf (accessed on 15 November 2019).
- Assis, F.; Fernando, L.; Ferreira, K.R.; Vinhas, L.; Maurano, L.; Almeida, C.; Carvalho, A.; Rodrigues, J.; Maciel, A.; Camargo, C. TerraBrasilis: A Spatial Data Analytics Infrastructure for Large-Scale Thematic Mapping. *ISPRS Int. J. -Geo-Inf.* **2019**, *8*, 513. [[CrossRef](#)]
- Beuchle, R.; Grecchi, R.C.; Shimabukuro, Y.E.; Seliger, R.; Eva, H.D.; Sano, E.; Achard, F. Land cover changes in the Brazilian Cerrado and Caatinga biomes from 1990 to 2010 based on a systematic remote sensing sampling approach. *Appl. Geogr.* **2015**, *58*, 116–127. [[CrossRef](#)]
- Bonanomi, J.; Tortato, F.R.; Raphael de Souza, R.G.; Penha, J.M.; Bueno, A.S.; Peres, C.A. Protecting forests at the expense of native grasslands: Land-use policy encourages open-habitat loss in the Brazilian cerrado biome. *Perspect. Ecol. Conserv.* **2019**, *17*, 26–31. [[CrossRef](#)]

19. Sathler, D.; Adamo, S.; Lima, E. Deforestation and local sustainable development in Brazilian Legal Amazonia: An exploratory analysis. *Ecol. Soc.* **2018**, *23*, 1–36. [[CrossRef](#)]
20. National Institute for Space Research (INPE). Detecting Residential Land-Use Development at the Urban Fringe. 1982. Available online: https://www.asprs.org/wp-content/uploads/pers/1982journal/apr/1982_apr_629-643.pdf (accessed on 13 October 2019).
21. Howarth, P.J.; Wickware, G.M. Procedures for change detection using Landsat digital data. *Int. J. Remote Sens.* **1981**, *2*, 277–291. [[CrossRef](#)]
22. Ludeke, A.; Maggio, R.C.; Reid, L. An analysis of anthropogenic deforestation using logistic regression and GIS. *J. Environ. Manag.* **1990**, *31*, 247–259. [[CrossRef](#)]
23. Nackaerts, K.; Vaesen, K.; Muys, B.; Coppin, P. Comparative performance of a modified change vector analysis in forest change detection. *Int. J. Remote Sens.* **2005**, *26*, 839–852. [[CrossRef](#)]
24. Deng, J.S.; Wang, K.; Deng, Y.H.; Qi, G.J. PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. *Int. J. Remote Sens.* **2008**, *29*, 4823–4838. [[CrossRef](#)]
25. Celik, T. Unsupervised change detection in satellite images using principal component analysis and *k*-means clustering. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 772–776. [[CrossRef](#)]
26. Xiaolu, S.; Bo, C. Change detection using change vector analysis from Landsat TM images in Wuhan. *Procedia Environ. Sci.* **2011**, *11*, 238–244. [[CrossRef](#)]
27. Zhan, Y.; Fu, K.; Yan, M.; Sun, X.; Wang, H.; Qiu, X. Change detection based on deep siamese convolutional network for optical aerial images. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1845–1849. [[CrossRef](#)]
28. Dhingra, S.; Kumar, D. A review of remotely sensed satellite image classification. *Int. J. Electr. Comput. Eng.* **2019**, *9*, 2088–8708. [[CrossRef](#)]
29. Kranjčić, N.; Medak, D.; Župan, R.; Rezo, M. Support Vector Machine Accuracy Assessment for Extracting Green Urban Areas in Towns. *Remote Sens.* **2019**, *11*, 655. [[CrossRef](#)]
30. Gunn, S.R. Support vector machines for classification and regression. *ISIS Tech. Rep.* **1998**, *14*, 5–16.
31. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [[CrossRef](#)]
32. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817. [[CrossRef](#)]
33. Valeriano, D.M.; Mello, E.M.K.; Moreira, J.C.; Shimabukuro, Y.E.; Duarte, V.; Souza, I.M.; dos Santos, J.R.; Barbosa, C.C.F.; de Souza, R.C.M. Monitoring tropical forest from space: The PRODES digital project. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *35*, 272–274.
34. National Institute for Space Research (INPE). Metodologia Utilizada nos Projetos Prodes e Deter. Available online: http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes/pdfs/Metodologia_Prodes_Deter_revisada.pdf (accessed on 15 October 2019).
35. Kintisch, E. Improved monitoring of rainforests helps pierce haze of deforestation. *Science* **2007**, *316*, 536–537. [[CrossRef](#)]
36. Shimabukuro, Y.; Duarte, V.; Anderson, L.; Valeriano, D.; Arai, E.; Freitas, R.; Rudorff, B.F.; Moreira, M. Near real time detection of deforestation in the Brazilian Amazon using MODIS imagery. *Ambiente Agua Interdiscip. J. Appl. Sci.* **2007**, *1*, 37–47. [[CrossRef](#)]
37. Brito, A.; Valeriano, D.D.M.; Ferri, C.; Scolastrici, A.; Sestini, M. Metodologia da detecção do desmatamento no bioma Cerrado. In *Mapeamento de Áreas Antropizadas Com Imagens de Média Resolução Espacial*; Instituto Nacional de Pesquisas Espaciais: São José dos Campos, Brazil, 2018.
38. Souza, C.; Azevedo, T. *MapBiomias General Handbook*; MapBiomias: São Paulo, Brazil, 2017; pp. 1–23.
39. Machado, R.B.; Ramos Neto, M.B.; Pereira, P.G.P.; Caldas, E.F.; Gonçalves, D.A.; Santos, N.S.; Tabor, K.; Steininger, M. *Estimativa de Perda da Área do Cerrado Brasileiro*; Relatório Técnico Não Publicado; Conservação Internacional: Brasília, Brazil, 2004; pp. 1–25.
40. Picoli, M.C.A.; Camara, G.; Sanches, I.; Simões, R.; Carvalho, A.; Maciel, A.; Coutinho, A.; Esquerdo, J.; Antunes, J.; Begotti, R.A.; et al. Big earth observation time series analysis for monitoring Brazilian agriculture. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 328–339. [[CrossRef](#)]
41. Zagoruyko, S.; Komodakis, N. Learning to compare image patches via convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 4353–4361.

42. Daudt, R.C.; Le Saux, B.; Boulch, A.; Gousseau, Y. Urban change detection for multispectral earth observation using convolutional neural networks. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 2115–2118.
43. Zhang, Z.; Vosselman, G.; Gerke, M.; Tuia, D.; Yang, M.Y. Change detection between multimodal remote sensing data using siamese CNN. *arXiv* **2018**, arXiv:1807.09562.
44. Mou, L.; Bruzzone, L.; Zhu, X. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 924–935. [[CrossRef](#)]
45. Lovett, G.M.; Burns, D.A.; Driscoll, C.T.; Jenkins, J.C.; Mitchell, M.J.; Rustad, L.; Shanley, J.B.; Likens, G.E.; Haeuber, R. Who needs environmental monitoring? *Front. Ecol. Environ.* **2007**, *5*, 253–260, doi:10.1890/1540-9295(2007)5[253:WNEM]2.0.CO;2. [[CrossRef](#)]
46. Rahman, F.; Vasu, B.; Van Cor, J.; Kerekes, J.; Savakis, A. Siamese Network with Multi-Level Features for Patch-Based Change Detection in Satellite Imagery. In Proceedings of the 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Anaheim, CA, USA, 26–29 November 2018; pp. 958–962.
47. Bazi, Y.; Melgani, F. Convolutional SVM networks for object detection in UAV imagery. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 3107–3118. [[CrossRef](#)]
48. Casseb, A.d.R.; Chiang, J.O.; Martins, L.C.; Silva, S.P.d.; Henriques, D.F.; Casseb, L.M.N.; Vasconcelos, P.F.d.C. Alphavirus serosurvey in domestic herbivores in Pará State, Brazilian Amazon. *Rev. Pan-Amazônica Saúde* **2012**, *3*, 43–48. [[CrossRef](#)]
49. Steinweg, T.; Gerard, R.; Thoumi, G. Cargill: Zero-Deforestation Approach Leaves Room for Land Clearing in Brazil's Maranhão. In Chain Reaction Research. 2018. pp. 1–18. Available online: <https://chainreactionresearch.com/wp-content/uploads/2018/04/Cargill-report-April-2018.pdf> (accessed on 20 November 2019).
50. Fan, R.E.; Chang, K.W.; Hsieh, C.J.; Wang, X.R.; Lin, C.J. LIBLINEAR: A library for large linear classification. *J. Mach. Learn. Res.* **2008**, *9*, 1871–1874.



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).