ASSESSMENT OF AN EARLY FUSION CNN APPROACH APPLIED TO THE DEFORESTATION DETECTION IN THE BRAZILIAN AMAZON

Mabel Ortega Adarme[1], *Patrick Nigri Happ*[1] *and Raul Queiroz Feitosa*[1,2]

[1] Pontifical Catholic University of Rio de Janeiro, Brazil - {mortega, patrick, raul}@ele.puc-rio.br
[2] Rio de Janeiro State University, Rio de Janeiro-RJ, Brazil

ABSTRACT

Deforestation is one of the main causes of biodiversity reduction, climate change among others destructive phenomena. Thus, early detection of deforestation processes is of paramount importance in the recent year. Motivated by this scenario the present work focuses on assessing a DL approach called Early Fusion (EF) for automatic deforestation detection. Change detection approaches based on Random Forest (RF) and Change Vector Analysis (CVA) were adopted as baselines for comparison purposes. These approaches were evaluated in a region located in the state of Pará, Brazil, where two images from Landsat 8 satellite were acquired to detect deforested areas from 2016 to 2017. Their corresponding references were collected from the Satellite Deforestation Monitoring Project in the Legal Amazon (PRODES). In the experiments, the EF approach outperformed RF and CVA baselines, identifying in a better way the regions that have suffered deforestation.

***Key words** – Deep learning, deforestation, image classification, early fusion, image stacking.*

1. INTRODUCTION

The Brazilian Amazon encompasses heterogeneous ecological and socioeconomic systems, which provide important environmental services not only to Brazil, but to the whole world [1]. Therefore, its preservation is essential due to a number of factors including the global ecological balance and climate change mitigation [2]. For decades, this system has been damaged by human activities, such as the legal and illegal felling of trees, extension of agricultural lands, construction of infrastructures and illegal mining.

In this context, the National Institute for Space Research (INPE), with the development of the Satellite Deforestation Monitoring Project in the Legal Amazon (PRODES) [3], supervises the deforestation in the Brazilian Legal Amazon (BLA) since 1988. Their objective is to quantify the deforestation of areas with native vegetation and, thus, to subsidize public policies to control and combat illegal deforestation. However, the adopted methodology involves a lot of manual operations. In this sense, an automatic detection could improve or at least alleviate the human hard working process.

Traditionally, change detection techniques based on image algebra, such as image differencing and Change Vector Analysis (CVA) [4], have been widely used to detect changes in multi-temporal images. Although not requiring prior information about the scene, they strongly depend on thresholds to define what is considered a change. In particular, CVA focuses on the analysis of differences to determine the changes in terms of strength and direction [5].

Recently, Deep Learning (DL) techniques have been successfully applied to Remote Sensing (RS) image analysis. Through the usage of Deep Neural Networks (DNNs), it is possible to learn multiple levels of data representation, which usually correspond to more informative features. In this respect, DNNs variants, such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) are potential candidates for automatic deforestation detection. Recently, Lyu et al. [4] proposed an end-to-end RNN to solve a multispectral image change detection problem, achieving encouraging results. Daudt et al. [6] successfully applied a CNN model to urban change detection based on an Early Fusion (EF) approach, which concatenates an image pair from different dates to build the input image for subsequent analysis.

The main objective of this paper is to adapt and to evaluate the aforementioned EF method for deforestation detection in Amazon region. Our study area is a region of BLA, located in the State of Pará, Brazil, where changes have been mapped from 2016 to 2017. For comparison purposes we take as baseline a Random Forest (RF) classifier and the CVA approach .

The remainder of this paper is organized as follows: Section 2 presents the change detection methods considered in this work. Section 3 describes the dataset and the experimental protocol. Results are presented in Section 4 and some concluding remarks pointing to future works are added in Section 5.

2. CHANGE DETECTION METHODS

2.1. Early Fusion CNN Stack (EF-stack)

CNN is a neural network, which involves a series of convolutional operations. Those operations use multiple kernel matrices to extract high-level features, which exploit spatial context among pixels. CNNs contain two basic operations: convolution and pooling, which are embedded in the sequential layers of the network [7]. Commonly, fully connected (FC) layers are added before the final classification is performed. For a deeper comprehension of CNNs, we refer to [8].

In this work, the CNN architecture model is inspired by [6], which achieved good results for urban changing detection. In this method, images of two dates, T1 and T2, are stacked by concatenating them along their spectral dimension. Then, patches of dimension w by w by $2c$ are extracted, where w is the spatial patch size and c is the number of spectral bands of each image. The CNN takes as input a single patch at time in order to assign a class label to the central pixel of each patch.

Galoá { Este trabalho foi publicado utilizando o Galoá proceedings

ISBN: 978-85-17-00097-3

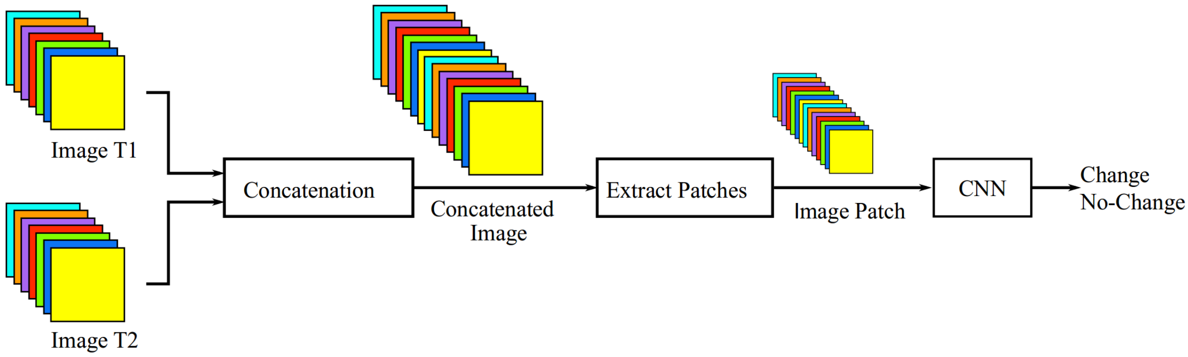


Figure 1: EF approach. Images at different dates (T1 and T2) are concatenated to form an image pair, then, patches are extracted and inputted to the CNN model.

This procedure is repeated for each image pixel as illustrated in (Figure 1).

2.2. Random Forest Stack (RF-stack)

RF is a supervised classification algorithm, which creates trees from randomly selected training samples and makes a prediction based on the majority vote of each output tree. In our analysis we follow a strategy similar to EF-stack. The inputs to the RF are the result of flattening a patch of size *w* by *w* by *2c* into a vector of dimension $w \times w \times 2c$. As in the previous approach, this procedure is carried out for each individual pixel location.

2.3. Change Vector Analysis (CVA)

CVA is an unsupervised technique in which magnitude and direction of changes are calculated. To determinate the change and no change areas, a threshold value might be selected. With the aim to detect the deforestation, the Normalized Difference Vegetation Index (NDVI) and the Bare Soil Index (BI) are calculated. NDVI quantifies vegetation by measuring the difference between near-infrared and red bands and BI is calculated to distinguish agricultural lands and non-agricultural lands [9]. These operations are executed for each pixel at dates T1 and T2. The NDVI and BI indices are calculated with Equations 1 and 2 respectively.

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

$$BI = \frac{(SWIR + RED) - (NIR + BLUE)}{(SWIR + RED) + (NIR + BLUE)} \quad (2)$$

where, *NIR*, *RED*, *SWIR* and *BLUE* are the spectral reflectance measurements acquired in the near-infrared, red, short wave infrared and blue regions.

The magnitude of the vector represents the changing intensity S and the direction represents the change type α. These values are obtained by Equations 3 and 4 respectively.

$$S = \sqrt{(NDVI_2 - NDVI_1)^2 + (BI_2 - B_I 1)^2} \quad (3)$$

$$tan(\alpha) = \frac{BI_2 - B_I 1}{NDVI_2 - NDVI_1} \quad (4)$$

$NDVI_1$, BI_1, $NDVI_2$, BI_2 represent the NDVI and BI indices computed on T1 and T2 respectively.

3. EXPERIMENTS

3.1. Study area

The study area is located in the State of Pará, Brazil, centered on coordinates of 03° 22' 12" S and 050° 42' 36" W. It corresponds to an area of the BLA, which has been facing a growing pressure of deforestation with a notable amount of deforested areas detected by PRODES [3]. The reference change map (Figure 2-a) is related to the deforestation occurred between 2016 and 2017 and was downloaded from PRODES database.

We also use a Landsat 8-OLI image pair, with 30m of spatial resolution, comprising the studied area. The images were acquired from the United States Geological Survey (USGS) in two different dates: August 2nd, 2016 (Figure 2-b) and July 20th, 2017 (Figure 2-c); with a size of 1300 $\times$ 1100 pixels and seven spectral bands (Coastal/Aerosol, Blue, Green, Red, NIR, SWIR-1, and SWIR-2) each. Additionally, an atmospheric correction was applied to each scene, and then, they were clipped to the area of interest. It is important to note that these dates were chosen due to the lower presence of clouds. In fact, PRODES reference were also generated using images from similar dates.

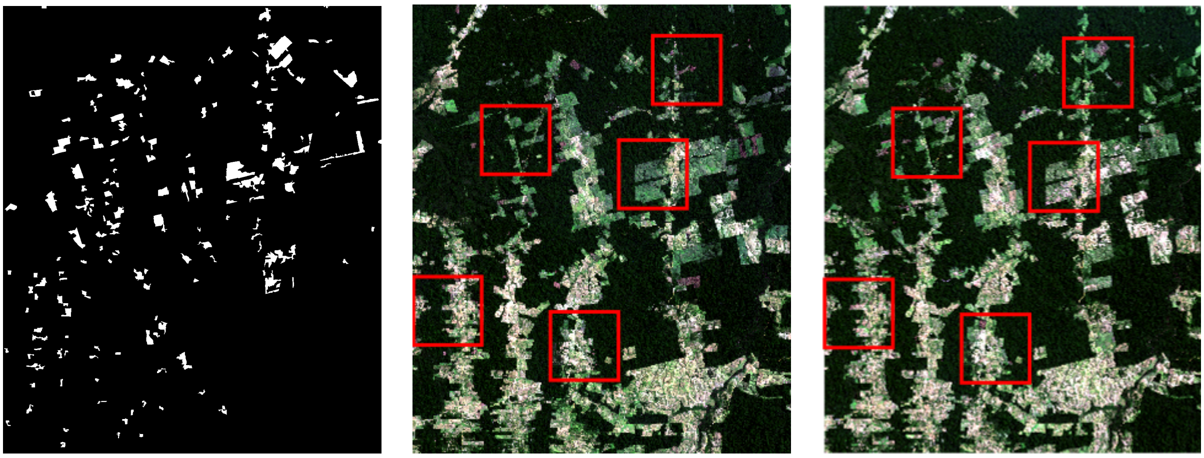
3.2. Experimental protocol

For the experiments with EF and RF methods, patches of size 15-by-15-by-14 were extracted from the stacked image pair considering all bands. This size was chosen after a preliminary experimental evaluation. Patches located inside the red squares shown in Figures 2 (b, c) were used to train the model, while the rest of them were used for testing. Each band was normalized to take values in the range -1 to 1.

We applied an under-sampling technique to balance the number of training samples for both classes (change and no-change). Thus, some randomly selected instances from the majority class (no-change) were removed to obtain an equal number of training samples (17,941 patches) for each class.

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

ISBN: 978-85-17-00097-3



(a) Ground truth (b) T1: 02/08/2016 (c) T2: 20/07/2017

Figure 2: Deforestation reference set (a) and RGB composition of the selected Amazon Forest region at dates T1 (b) and T2 (b). Red boxes indicate the training set.

Layer	Filter Size	Output Size	Parameters
Input	-	15 ×15 × 14	0
Conv1	4 ×4	15 ×15 × 96	21600
Conv2	3 ×3	15 ×15 × 96	83040
Conv3	3 ×3	15 ×15 × 96	83040
Conv4	3 ×3	15 ×15 × 96	83040
MaxPool	2 ×2	7 ×7 × 96	0
Conv5	3 ×3	7 ×7 × 192	166000
Conv6	3 ×3	7 ×7 × 192	331968
Conv7	3 ×3	7 ×7 × 192	331968
Flatten	-	1 ×9408	0
Dense	-	1 ×2	18818

Table 1: Parameters of the EF architecture.

We considered four metrics to evaluate the performance of the three methods: Precision, Recall, F1-score and Overall Accuracy (OA).

The CNN architecture used in the EF approach is composed of seven convolutional layers (Conv) with ReLU activation function, a max-pooling layer (MaxPool), a Flatten layer and a Dense layer with a Softmax activation function for the two outputs, associated with the classes deforestation (change) and no deforestation (no-change). This architecture is summarized in Table 1. The EF model was trained with 100 epochs, batch size of 64 and with the Adam optimizer, in contrast to the [6], which used Average Stochastic Gradient Descent (ASGD). An early stop was added as regularization technique to avoid over-fitting.

For the RF classifier, the number of random trees and its maximum depth was set to 300 and 30, respectively. For CVA, a pixel-wise approach was employed for the estimation of NDVI and BI indices. We considered only the magnitude of change, because in our setup there is just one type of change, namely whether or not the target area underwent a deforestation. The change map is generated using a threshold of 0.6, which was selected empirically.

4. RESULTS

Table 2 summarizes the results obtained in terms of Precision, Recall, F1-score and Overall Accuracy (OA) values for each method described in Section 2. Note that the EF method achieved the best results in terms of three out of the four metrics, identifying more accurately the the deforestation. This is probably due to the ability of CNNs to learn more informative features and to model non-linear and complex input-to-output relationships. In terms of Precision, RF was the best performing strategy for the class deforestation (change). However, the Recall was lower, indicating that there RF tends to produce more false no deforestation (no-change) outcomes than the other methods. .

CVA presented the lowest success rate for deforestation detection and the worse result in terms of OA. The poor results of this method may be related to the indices adopted to compute the changing intensity, especially the BI. Additionally, CVA is strongly conditioned to the threshold definition and the threshold determination depends on the skills of the analyst. On the other hand, it presented very good results for the class no deforestation (no-change) .

The Overall Accuracy achieved by EF and RF were greater than 90%. However, the results related to the deforestation detection were lower than 60% in terms of F1-Score. This might be due to the different causes of changes, such as agricultural or pastoral farming and mining, which leave distinct footprints in the image.

Figure 3 shows a snip of the change maps delivered by each method, where true change (yellow), false change (red), false no-change (blue) and true no-change (white) are represented. We considered as true change the deforestation pixels correctly identified and true no-change the no deforestation pixels that were classified as it. Taking these images into account, it is possible to perceive how the EF method better identified deforestation with low false no-change (Figure 3-a) in contrast with RF, which generated a greater number of false detection as can be appreciated in Figure 3-b. Likewise, as it can be observed in Figure 3-c, the deforestation detection was worse for CVA. In this case, there are a lot of samples that were not identified as deforestation regions and a salt and pepper effect can be observed.

Galoá { Este trabalho foi publicado utilizando o Galoá proceedings

ISBN: 978-85-17-00097-3
INPE - Santos-SP, Brasil

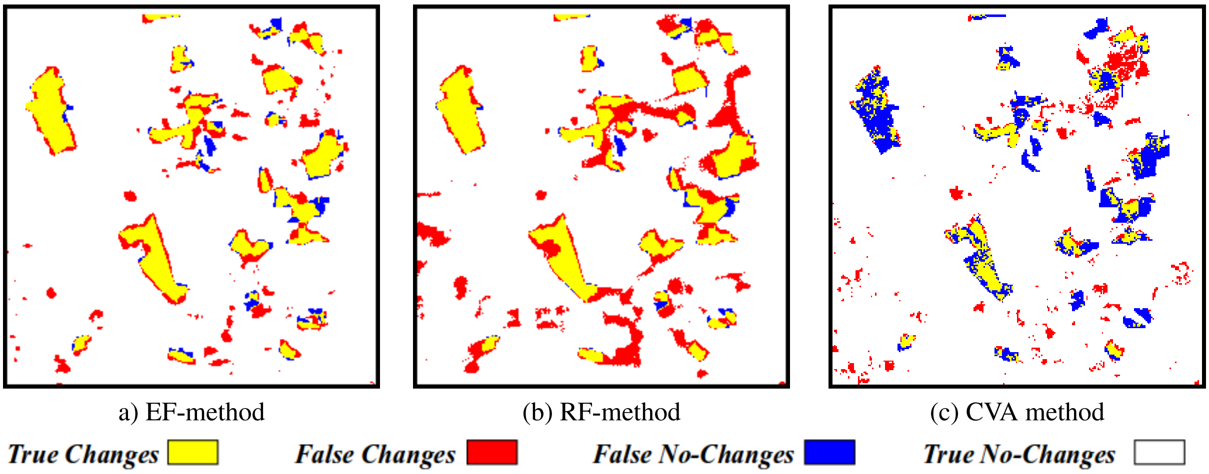


Figure 3: Comparison of the change map obtained from the EF, RF and CVA methods on a small scene of the test region.

Metric (%)	Method	Change	No Change
Precision	EF	65.67	97.04
	RF	**78.22**	93.85
	CVA	35.10	97.89
Recall	**EF**	**49.95**	97.97
	RF	41.12	96.49
	CVA	29.25	98.38
F1-Score	**EF**	**56.74**	97.50
	RF	53.90	95.15
	CVA	31.91	98.14
OA	**EF**	**96.11**	
	RF	93.98	
	CVA	82.78	

Table 2: Metrics values obtained from EF, RF and CVA methods.

5. CONCLUSIONS

In this work, an evaluation of an EF CNN method for deforestation detection in a region of the BLA was reported. The analysis revealed that the EF model overcomes the traditional RF and CVA approaches.

EF, being a model based on a deep learning architecture, presents advantages such as automatic feature learning and flexibility, which allows improving the representation of the relationships between the two images (T1 and T2), therefore improving the deforestation detection performance.

In terms of OA, the results obtained by the EF are particularly good. However, the detection rate for deforestation, in terms of F1-score, was lower than 60%. This shows the complexity and the difficulty of this kind of application. Thus, although the results are still not good enough, this work can be considered as an initial study and a step further to possible solutions to improve the automatic deforestation detection.

Future works are intended to explore different architectures as well as changes on the adopted configuration, like using as input the values from the artificial vegetation, soil and shadow bands that are used on the PRODES methodology. Another investigation is related to the usage of Synthetic Aperture Radar (SAR) data to complement the information already available from the optical sensor.

ACKNOWLEDGEMENTS

The authors acknowledge the funding provided by CAPES and CNPq.

6. REFERENCES

[1] SOUZA, R. A. de; JUNIOR, P. D. M. Improved spatial model for amazonian deforestation: An empirical assessment and spatial bias analysis. *Ecological Modelling*, Elsevier, v. 387, p. 1–9, 2018.

[2] SATHLER, D.; ADAMO, S.; LIMA, E. Deforestation and local sustainable development in brazilian legal amazonia: an exploratory analysis. *Ecology and Society*, The Resilience Alliance, v. 23, n. 2, 2018.

[3] (INPE), N. I. for S. R. *Monitoring of the Brazilian Amazonian Forest by Satellite*. 1988. Avaliable at: <http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>.

[4] LYU, H.; LU, H.; MOU, L. Learning a transferable change rule from a recurrent neural network for land cover change detection. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 8, n. 6, p. 506, 2016.

[5] XIAOLU, S.; BO, C. Change detection using change vector analysis from landsat tm images in wuhan. *Procedia Environmental Sciences*, Elsevier, v. 11, p. 238–244, 2011.

[6] DAUDT, R. C. et al. Urban change detection for multispectral earth observation using convolutional neural networks. In: *International Geoscience and Remote Sensing Symposium (IGARSS)*. [S.l.: s.n.], 2018.

[7] BUDAK, Ü.; ŞENGÜR, A.; HALICI, U. Deep convolutional neural networks for airport detection in remote sensing images. In: IEEE. *2018 26th Signal Processing and Communications Applications Conference (SIU)*. [S.l.], 2018. p. 1–4.

[8] GOODFELLOW, I. et al. *Deep learning*. [S.l.]: MIT press Cambridge, 2016. v. 1.

[9] SI, S.; THI, L.; VAN, C. Land cover change analysis using change vector analysis method in duy tien district, ha nam province in vietnam. In: *7th FIG Regional Conference, Hanoi*. [S.l.: s.n.], 2009. p. 19–22.

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings