

MAPPING EUCALYPTUS PLANTATIONS AND NATURAL FOREST AREAS IN LANDSAT-TM IMAGES USING DEEP LEARNING

Matheus Pinheiro Ferreira¹, Laura Elena Cué La Rosa², Patrick Nigri Happ², Raissa Brand Theobald¹ and Raul Queiroz Feitosa^{2,3}

¹ Cartography Engineering Section, Military Institute of Engineering, Rio de Janeiro-RJ, Brazil – matheus@ime.eb.br

² Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro-RJ, Brazil – (lauracue, patrick, raul)@ele.puc-rio.br

³ Rio de Janeiro State University, Rio de Janeiro-RJ, Brazil

ABSTRACT

Automatic mapping of planted and natural forests using satellite images is a challenging task due to spectral similarity issues. In this work, we assessed the use of Convolutional Neural Networks (CNNs) to discriminate between natural forest areas and eucalyptus plantations in a Landsat-TM scene. First, we produced training and testing datasets with data from the MapBiomass project. Then, CNNs were trained with input patches of different sizes (5×5 , 7×7 , 9×9 and 11×11 pixels) to evaluate the influence of patch dimension in the classification accuracy. For comparison, pixel-wise and patch-classification were performed using the Random Forest (RF) algorithm. The best results were obtained using CNNs with 5×5 patches. In this scenario, the F-score was of 97.64% for natural forests and 95.49% for eucalyptus plantations. The classification errors reached 9.06% using RF and did not exceed 3% with CNNs.

Key words – Convolutional Neural Networks, patch-classification, random forest, satellite images, tropical forests

1. INTRODUCTION

Since 1990, the global extent of planted forests increased 65% due to the growing demand for forest and timber products [1]. In Brazil, the area of planted forests increased some 47 million ha from 1985 to 2017 [2]. The most planted woody genus in Brazil is *Eucalyptus* spp., occupying an area of about 5.7 million ha [3]. At the same time, natural forests are facing deforestation and degradation, with the tropics alone accounting for 32% of the global forest loss from 2000 to 2012 [4].

Remote sensing has been widely used to map and monitor forested areas at a wide range of temporal and spatial scales. Particularly Landsat images have been successfully employed to produce national and regional forest cover maps. However, most of the mapping initiatives do not discriminate between natural and plantation forest areas [4] or, if they do, a high degree of misclassification is present [2]. Mapping tree plantations and natural forests in satellite images, such as those from Landsat, is challenging due to their spectral similarity.

In areas of old occupation of Brazil, in which agricultural or urban fields replaced the original forest cover, natural forests are characterized by small (< 100 ha) and disconnected fragments located mainly along river beds [5]. Conversely, eucalyptus plantations are usually large areas (> 500 ha), forming homogeneous and monospecific

patches. Both natural forest areas and eucalyptus plantations have distinct spatial distribution patterns. Thus, machine learning methods that take into account the spatial context can improve discrimination between these classes. Convolutional Neural Networks (CNNs), a type of deep learning method, use spatial-contextual information for classification and has been producing excellent results for remote sensing image classification.

In this work, we assessed CNNs to discriminate between natural forest areas and eucalyptus plantations in a Landsat Thematic Mapper (TM) scene. Training and testing datasets were produced by using data from the MapBiomass project [2]. We ran CNNs with different patch configurations to evaluate the influence of patch size in the classification accuracy. For comparison purposes, we also performed patch-classification using the Random Forest (RF) algorithm.

The remainder of this paper is organized as follows: Section 2 presents the study area and the forest cover data. Section 3 describes the methods adopted as well as the experimental set-up. Section 4 discuss the experimental results and finally, Section 5 presents the conclusions.

2. MATERIAL

2.1. Study site and remote sensing data

For this study, we selected the Landsat-TM worldwide reference system 2 (WRS2) 220/76 scene (Figure 1a) that was acquired on 14 April 2006 over the center-south portion of the São Paulo state, southeastern Brazil. The region shows highly diverse land use patterns characterized by silvicultural (eucalyptus plantations) and agricultural (mainly sugarcane) fields. Natural forest areas are sparsely distributed in small (< 100 ha) fragments that are primarily in riparian areas [6]. The Landsat-TM image was retrieved from the global visualization viewer (GloVis) (<https://glovis.usgs.gov/>) in surface reflectance (L2 level).

2.2. Forest cover data

Aiming to assess CNNs for mapping eucalyptus plantations and natural forest areas, we used the land cover dataset produced by the MapBiomass project. MapBiomass employs a pixel-wise classification scheme, based on the U.S Geological Survey Landsat Global Archive, to regularly map more than 20 land use/cover classes over the Brazilian territory. More details about the project can be found at <http://mapbiomas.org>. We used the available 3.0 collection and selected the class IDs that represented natural forest areas (IDs=1 to 5) and eucalyptus plantations (ID=9). The pixels

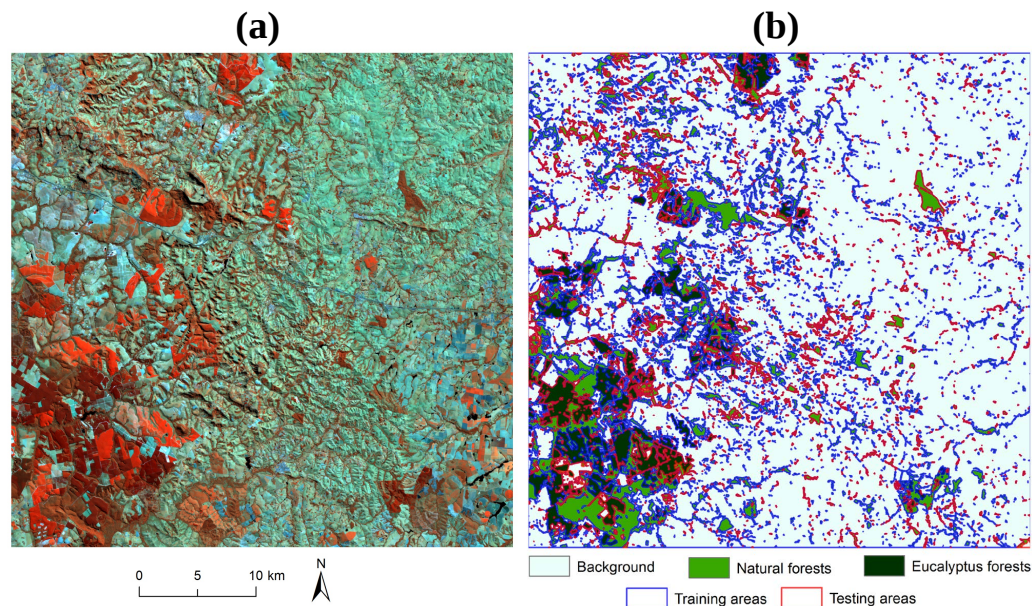


Figure 1: (a) false color composition (RGB 453) of the Landsat-TM image used in this work. (b) forest cover map showing natural forests and eucalyptus plantations retrieved from the MapBiomass project. Training and testing areas are depicted in blue and red, respectively.

from the other classes were set to zero and assigned to a class called background (Figure 1b).

3. METHODS

CNNs are one of the most used deep learning methods for remote sensing image classification and have been achieving outstanding performance levels [7]. CNNs constitute a class of deep Artificial Neural Network which rely on local linear operations (convolutions) followed by non-linear transformations, creating different representations of the input data.

The convolutional layers act as feature extractors from the input images. These layers are composed of a set of filters that encode lower-level features (from the first layers) into more high-level features (from deeper layers) taking into account the spatial context. Generally, a non-linear activation function is applied to the output of a convolutional layer, which is followed by a downsampling (pooling) process to reduce its dimensionality. After several convolutional and pooling layers, a fully-connected (FC) layer might be included to exploit the high-level features learned, which could be seen as hidden layers of a Multilayer Perceptron (MLP). The last layer of a CNN is often a *softmax* classifier that outputs class membership probabilities for each class. In addition, several state-of-the-art architectures use Batch Normalization [8] to make the network training process less sensitive to layer initialization and also to improve its convergence. This is done by forcing the set of activations of the previous layer to have zero mean and unit variance. For a comprehensive overview of CNNs and deep learning, the reader is referred to [9].

As in [10], we adopt the CNNs patch-classification approach. This approach captures the spatial context taking as input an image patch (extracted from the original image) and predicts a single label, which is assigned to the central pixel of

the patch. As illustrated in Table 1 the CNN architecture used in this work is constructed by 3 convolutional (Conv) layers and 2 pooling (Pool) layers. After each convolution, a batch normalization is applied as well as a non-linear activation function (BnAct) to get the output feature map.

Type	Filter Size/ Stride	Output Size	Params
Conv1	$3 \times 3/1$	$9 \times 9 \times 32$	1760
BnAct1	-	$9 \times 9 \times 32$	128
Pool1	$2 \times 2/2$	$5 \times 5 \times 32$	-
Conv2	$1 \times 1/1$	$5 \times 5 \times 48$	1584
BnAct2	-	$5 \times 5 \times 48$	192
Conv2	$3 \times 3/1$	$5 \times 5 \times 64$	27712
BnAct2	-	$5 \times 5 \times 64$	256
Pool2	$2 \times 2/2$	$3 \times 3 \times 64$	-
FC1	-	128	73856
BnAct3	-	128	512
Softmax	-	2	258
Total	-	-	106.2K

Table 1: Architecture of the CNNs model. Example case for input image patch of $9 \times 9 \times 6$.

For the sake of comparison, we also performed patch-classification using Random Forest (RF) [11], a very common machine learning approach for image classification. Similarly to the CNNs architecture described above, only the central pixel of each image patch was labeled. For both methods, the procedure consists of three main steps: (i) the input image is cropped in densely overlapping image patches with a sliding window technique with stride 1 to preserve the spatial resolution, (ii) a CNN/RF method is used to perform both training and inference and (iii) each patch is spatially concatenated to obtain a classification map at the same resolution of the input image. It is worth noting that, for the RF classifier, each image patch is flattened and act as the feature vector correspondent to the central pixel of the patch.

	1x1		5x5		7x7		9x9		11x11	
	NF	EP	NF	EP	NF	EP	NF	EP	NF	EP
CNNs	-	-	97.64	95.49	95.61	91.14	97.06	94.27	96.92	93.92
RF	91.18	79.80	95.05	89.38	94.76	88.64	94.30	87.45	93.87	86.30

Table 2: F-score obtained using CNNs and RF to map natural forest (NF) areas and eucalyptus plantations (EP) with different patch sizes. The highest F-scores values are highlighted in bold.

3.1. Experimental set-up

For the classification experiments, the forest cover map was divided into training and testing sets. First, a raster-to-vector conversion was performed to obtain polygons of natural forests and eucalyptus plantations. Then, these polygons were randomly partitioned into 70% for training and 30% for testing (Figure 1b).

We are interested to know how patch size influences the classification results of CNNs and RF. Thus, we used five patch dimensions in our experiments: 1×1 , 5×5 , 7×7 , 9×9 and 11×11 pixels. To assess the final classification maps we computed the F-score, which is a weighted average of the precision and recall. It is important to highlight that only two classes were considered in this study (natural forest areas and eucalyptus plantations); the background pixels were set to zero and did not influence the accuracy assessment.

The hyperparameters of the CNNs method were tuned based on experiments. In order to increment the number of training samples, we applied a data-augmentation procedure (rotation and flips). The batch size was selected experimentally and fixed to 128. For the optimization, we used Adam optimizer [12] with a learning rate of 0.001. As non-linear transformation, we selected the Leaky version of a Rectified Linear Unit (ReLU) function [13]. The RF method was implemented using the Sklearn module of Python and Keras with TensorFlow backend for the CNNs. The models were trained on a desktop workstation with an Intel Core i7-4790 3.6GHz CPU, 32GB of main memory and an NVIDIA GeForce GTX1080 graphics processor with 12GB of memory. All experiments ran under Linux (Ubuntu 16.04 distribution).

4. RESULTS AND DISCUSSION

In Table 2 we show the F-scores obtained for CNNs and RF with different patch dimensions. Patches of 1×1 pixel represent the pixel-wise classification approach, thus neglecting the spatial context. The best results were obtained by using 5×5 patches. For classifications with RF, the adoption of such patches showed to increase the F-score by up to 3.87% for natural forests and 9.58% for eucalyptus plantations, in comparison to the pixel-wise approach. However, patches larger than 5×5 pixels decreased the F-scores of RF for the two classes considered. For CNNs, the best results were also observed by adopting 5×5 patches (Table 2). Conversely to the results of RF, the lowest F-scores were not observed with the largest patches (11×11), but with those of 7×7 pixels. More experiments are needed to evaluate the reasons why this patch dimension provided the poorest results with CNNs.

Figure 2 shows the classification accuracy (diagonal cells)

CNNs			RF		
True class	Predicted class		True Class	Predicted class	
	NF	EP		NF	EP
NF	96,87	3,03	NF	90,94	1,00
EP	3,13	96,97	EP	9,06	99,00

Figure 2: Confusion matrices showing the classification accuracy of Natural Forests (NF) and Eucalyptus Plantations (EP) obtained using CNNs and RF with patches of 5x5 pixels.

and classification errors (off-diagonal cells) of each class for maps produced by CNNs (a) and RF (b) with 5×5 patches. Both classifiers achieved an accuracy greater than 90% on the testing set, with CNNs yielding the highest values. The classification errors between natural forests and eucalyptus plantations were of about 3% for CNNs and up to 9% for RF. The error distribution between the classes was more balanced for CNNs. The RF tended to classify more eucalyptus plantation pixels as natural forests.

Spatial subsets of the maps produced by CNNs and RF with different patch dimensions are shown in Figure 3. The red circle highlights an area of eucalyptus plantation. We can note that within this area, the number of pixels labeled as natural forests increases according to the dimensions of the patch used for classification. Apparently, smaller patches (5×5), are able to better characterize spatial-contextual information of eucalyptus plantations. Additional experiments are needed to better understand the influence of patch size.

5. CONCLUSIONS

In this work, we evaluated the use of CNNs, a deep learning model, to map natural forest areas and eucalyptus plantations in southeastern Brazil. Data from the MapBiomass project was used to train CNNs and a classification error of 3% was obtained. CNNs also outperformed RF to discriminate between natural forests and eucalyptus plantations. Experiments performed with patches of different sizes showed that smaller patches (5×5) provided the best classification results. Future work will incorporate textural features in the classification process and will explore other proportions of training and testing samples.

ACKNOWLEDGEMENTS

The authors acknowledge the funding provided by CAPES and CNPq.

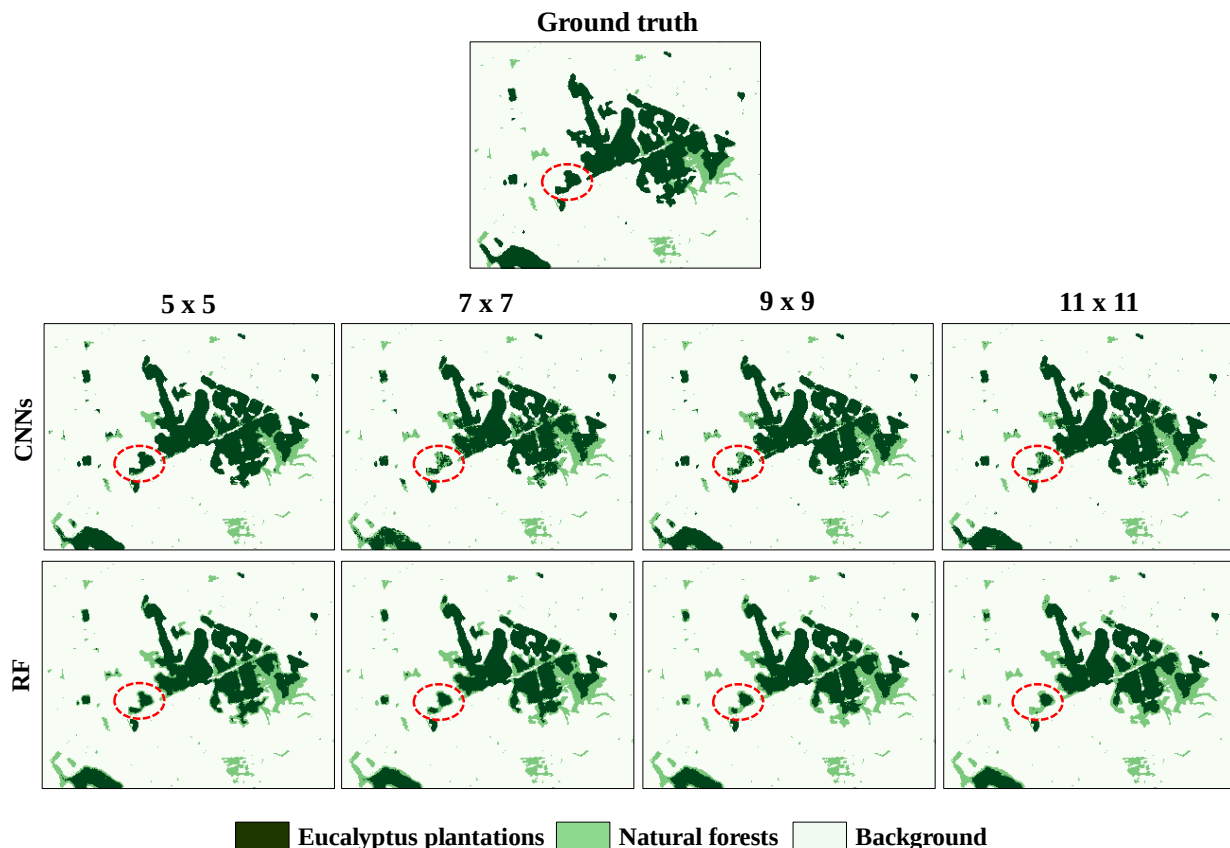


Figure 3: Classifications maps produced with different patch sizes using convolutional neural networks (CNNs) and Random Forest (RF). The red circle highlights an area of eucalyptus plantation.

6. REFERENCES

- [1] FAO. *Global Forest Resources Assessment 2015*. Food and Agriculture Organization of the United Nations. Rome, Italy, 2015.
- [2] MAPBIOMAS. *Collection 3.0 of the annual series of land use and cover of Brazil*, accessed in 02 Oct. 2018 at <http://mapbiomas.org>. [S.l.], 2017.
- [3] IBA. *Brazilian Tree Industry report 2017*. São Paulo, SP, Brazil, 2017.
- [4] HANSEN, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science*, American Association for the Advancement of Science, v. 342, n. 6160, p. 850–853, 2013.
- [5] RIBEIRO, M. C. et al. The brazilian atlantic forest: How much is left, and how is the remaining forest distributed? implications for conservation. *Biological conservation*, Elsevier, v. 142, n. 6, p. 1141–1153, 2009.
- [6] SILVA, A. L.; ALVES, D. S.; FERREIRA, M. P. Landsat-based land use change assessment in the brazilian atlantic forest: Forest transition and sugarcane expansion. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 10, n. 7, p. 996, 2018.
- [7] ZHANG, L.; ZHANG, L.; DU, B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, IEEE, v. 4, n. 2, p. 22–40, 2016.
- [8] SHIMODAIRA, H. Improving predictive inference under covariate shift by weighting the log-likelihood function. *Journal of statistical planning and inference*, Elsevier, v. 90, n. 2, p. 227–244, 2000.
- [9] PONTI, M. A. et al. Everything you wanted to know about deep learning for computer vision but were afraid to ask. In: *IEEE. Graphics, Patterns and Images Tutorials (SIBGRAPI-T)*, 2017 30th SIBGRAPI Conference on. [S.l.], 2017. p. 17–41.
- [10] VOLPI, M.; TUIA, D. Dense semantic labeling of subdecimeter resolution images with convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, v. 55, n. 2, p. 881–893, 2017.
- [11] BREIMAN, L. Random forests. *Machine learning*, Springer, v. 45, n. 1, p. 5–32, 2001.
- [12] KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [13] MAAS, A. L.; HANNUN, A. Y.; NG, A. Y. Rectifier nonlinearities improve neural network acoustic models. In: *Proc. icml*. [S.l.: s.n.], 2013. v. 30, n. 1, p. 3.