A NOVEL APPROACH TO RECOGNIZING PATTERNS IN REMOTE SENSING TIME-SERIES USING DEEP LEARNING

Evandro Taquary[1], Leandro Leal Parente[2], Ana Paula Matos e Silva[3], Laerte Guimarães Ferreira[4]

[1]evandro.taquary@gmail.com; [2]leal.parente@gmail.com; [3]annapaulamatos1@gmail.com; [4]laerte@ufg.br;
Image Processing and GIS Lab at the Federal University of Goiás (LAPIG/UFG),
CAMPUS II Samambaia, Goiânia – GO, Brazil

ABSTRACT

One of the most remarkable breakthroughs of Remote Sensing lies upon the devise of CubeSat standard. Such technology open up a myriad of possible applications that benefit from the higher spatio-temporal resolutions delivered by constellations of CubeSat compliant nanosatellites. Within this scenario, one has to investigate the new challenges and how to tackle them in order to harness this new kind of Remote Sensing Big Data. Among these challenges is the development of the means to extract useful information of pixels' observations throughout time in a fine-grained fashion. This work is a seminal study on using a special kind of *deep learning* approach, namely, deep Recurrent Neural Networks, for classifying long time-series of landcover's observations. The method was tested against the problem of identifying pastureland areas over high-res imagery from PlaneScope, a constellation of CubeSat nanosatellites. A discussion concerning limitations and capabilities of the proposed approach are also presented.

Key words — deep learning, recurrent neural networks, high resolution imagery, planet images, pasturelands

1. INTRODUCTION

The increasing number of sensors orbiting the earth is systematically producing larger volumes of data, with better resolutions in both the spatial and temporal dimensions [1]. This trend can be noted by the rise of the so called CubeSat constellations, in which many nanosatellites, a.k.a nanosat, are deployed to operate in concert, sharing the same goal: to provide better land surface coverage, as well as higher revisit frequency [2].

In the wake of the great cost benefit brought by CubeSat constellations, Planet Labs Inc. deployed more than 130 cheap CubeSats [3], known as PlaneScope constellation, which has a near daily basis revisit rate at nadir and high spatial resolution of around 3~5m. Although the unprecedented improvements of spatio-temporal resolutions, the main downsides of nanosat technologies are the poor radiometric quality and cross-sensor inconsistencies [3].

To deal with the high degree of details provided by CubeSat constellations, as well as with the low radiometric quality, more accurate machine learning techniques, such as *deep learning* (DL), are required to transform raw data into useful information [4]. Concerning spatial resolution, some DL architectures (i.e. Convolutional Neural Networks and its variants) rely upon only spatial dimensions to perform, for example, land-cover/land-use (LCLU) maps, disregarding the temporal dependencies between pixels observations over time [5].

Remote Sensing (RS) data with high temporal resolution (*e.g.* PlanetScope, Sentinel 2) may provide more consistent time-series that can be used, for example, to identify important LCLU classes like crop, pastureland, and grasslands. Due to seasonality and the phenological aspects that characterize these targets, an approach able to recognize patterns in their temporal signatures benefits from the fine-grained temporal detailing.

This potential can be explored using deep Recurrent Neural Networks (RNN), a specific family of DL approaches able to take all available observations throughout time into account. This way, latent information that lies between all the pixel's observations become intelligible, making possible to extract it from time-series directly and boosting a model's capability of recognizing temporal patterns.

This work presents an implementation of a specialized kind of RNN called Long Short Term Memory (LSTM) to classify pixels' time-series from Planet Labs imagery in order to identify pasturelands. It is a first trial, aimed to validate a novel approach to extracting temporal information from high resolution imagery through DL algorithms. In this sense, more improvements has to be taken for the approach to become operational. The model's accuracy and spatial consistency of the produced map were evaluated and a discussion concerning limitations and capabilities of the proposed methods are also provided.

2. DATA AND METHODS

In order to investigate the feasibility of using deep RNNs to classify SR time-series, a pastureland mapping has been performed. As a proof of concept, the mapping considered a reduced area comprising the Itapirapuã brazilian municipality, a Cerrado region located in the center-west of the state of Goiás, in which cattle ranching is the main economic activity. This choice was motivated assuming the premise that the phenological patterns of pasturelands can be easily recognized by a human interpreter as well as a deep RNN, as the seasonality of the chosen region is easily recognizable due to presence of dry and wet seasons.

The mapping has been performed using high-res PlanetScope imaging in which each pixel covers approximately a land surface area of 9 (3x3) m^2, and has a revisiting rate of about 1 to 2 days. Images dating between 01/08/2017 and 31/12/2017 that intercept the region of interest were acquired through Planet Explorer using the following filters: less than 1% of cloud cover; area coverage (considering all images within a day) greater than 99%; and, data source equals to "4-band PlanetScope Scene". All images that composes the daily scenes that meet the filters criteria were downloaded and processed to create 52 daily compositions with four spectral bands (i.e. blue, green, red, and near infrared). It is worth mentioning that the availability distribution of images throughout time wasn't uniform, since the wet season begins around October and cloud cover is higher within this period. Also, many days didn't meet the “area coverage” criteria for the study region and were then neglected.

The proposed approach required a set of training data compatible with the spatial resolution of PlanetScope images. In this direction, a protocol for picking up training points and its corresponding time-series was developed. The first thing to be considered in this scenario is to increase the number of training samples in order to better represent the actual distribution of the pixels' time-series within the feature space regarding the region of interest. The daily compositions were used to select, by visual inspection, a set of sample points, shown in figure 1, as described below:

1. Using a Landsat based reference map, 2300 points were randomly drawn over pasture areas and 2300 points over non-pasture areas;
2. The first stage of visual inspection, using PlanetScope images and / or Google Earth, corrected badly placed points of pasture and non-pasture classes (due to the use of a reference map with coarser spatial resolution, i.e. 30m);
3. The second visual inspection compared the first and last daily composition aiming to move the points to areas without LCLU change in the period;
4. The third visual inspection balanced the training points between pasture and non-pasture subclasses, such that more prevalent subclasses points were moved to less incident subclasses.

Figure 1. Sample set, grazing points in green and not pasture in yellow, used to train and evaluate the LSTM model in the identification of pasture areas, around the municipality of Itapirapuã - Goiás.

The sample set was divided into two subsets: training (70%) and test (30%) points, which were crossed with the 52 daily compositions to extract the 52 length labeled time-series. The elaboration of the model took place through several executions and performance metrics evaluations (accuracy and loss), as well as visual analysis of the classified pasture area. The RNN model was implemented using the Keras library and progressively refined through continuous training samples adjustments until reaching the following architecture and hyperparameters:

- Architectural layers:
  - A LSTM cell with 256 hidden units;
  - A fully-connected output layer with a softmax activation function;
- Loss function: Categorical Cross-Entropy
- Optimizer: Adam
- Learning Rate: 0.00002
- Batch size: 16
- Number of Epochs: 200

All the chosen hyperparameters and network's topology were discovered empirically, i.e, there's no known way to find out these technical features *a priori*, so that there must be performed a number of sensitivity analysis rounds in order to reach out the optimal set of hyperparameters and model's architecture.

Tests were performed within a Docker container with a Keras-ready environment set up, running on a host equipped with 12 CPUs Intel(R) Xeon(R) CPU E5-2620 v2 @ 2.10GHz and 24GB of RAM. No graphic cards (i.e. GPGPU) were used to haste execution of training nor prediction phases. The source code used by this study is available at https://github.com/NexGenMap/dl-time-series.

3. RESULTS AND DISCUSSIONS

The training behavior of the LSTM presented similar loss and accuracy curves for training and test datasets (figure 2), showing little shift between them and indicating absence of overfitting, i.e., a good model's generalization ability. The loss curve shows values converging toward approximately 0.2, suggesting that model's classification efficiency can be improved through modifications in the network architecture, such as the inclusion of new architectural layers and/or the use of regularization techniques (e.g dropout), such that this value approaches 0.

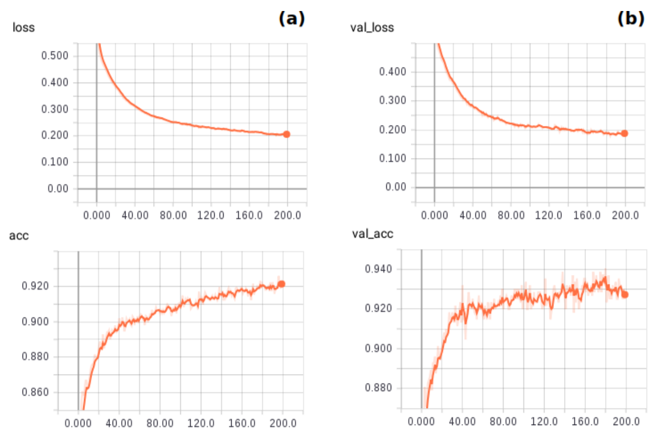

Figura 2. Loss and accuracy curves, produced during the LSTM training process, for (a) training data and (b) test data.

The mapping of pasture areas with LSTM showed a good spatial consistency and good separability between different land use classes (figure 3), such as tree plantation, bare soil, river, lake, natural vegetation, road and urban area (figure 3a, 3b and 3d). Specifically in the urban area, only grassland and square areas with presence of grass were classified as pasture, indicating a potential use of this approach, and PlanetScope imagery, for measuring the proportion of green area in urban densities. In pasture areas, the mapping generally included only pasture pixel with presence of shrubby vegetation, maintaining pixels with presence of arboreal vegetation as "non-pasture" (Figure 3c).

A significant advantage of LSTM - and other DL algorithms - in relation to Random Forest and other classical machine learning algorithms, is the absence of the feature engineering stage. This allowed the use of complete time-series, with 52 daily compositions, to find the most suitable seasonal patterns for pasture classification, directly in the 4 spectral bands. Beyond its ability of dealing with long time-series, a LSTM can also identify seasonal patterns by even learning new implicit “spectral indices” that are more efficient than, for example, handcrafted spectro-temporal metrics that make use of NDVI to separate pastures from other LCLU classes, like other known approaches does.

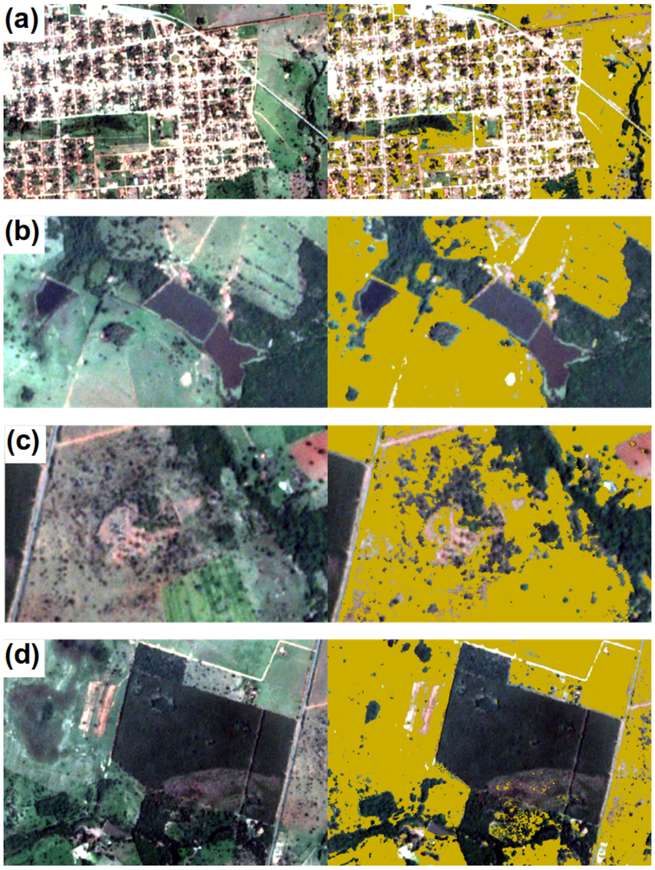

Figure 3. Result of the mapping of pastures, produced by the LSTM, in regions of (a) urban area; (b) lakes and natural vegetation; (c) foul grass; (d) forestry and unpaved roads. Interestingly, in the areas of foul grass, most of the tree vegetation was not classified as pasture.

Despite the promising results, the current implementation has some limitations with regarding spatio-temporal generalization. The model was conceived as a proof of concept, so it is very fitted to the region and period considered in the experiment. Not all days that comprise the chosen period (August to December of 2017) satisfy the restraints imposed by the aforementioned filters; this incur in a very unbalanced distribution amongst days that comply with the constraints and others that do not, what makes time-series not chronologically uniform and very specific to the region of interest, given a span of time.

Nothing prevents one to create as much models as needed to comprehend a wider region, using the exact same implemented approach, but another way to overcome this issue is to take timesteps (i.e. daily compositions) with presence of nodata values into account by, for example, loosening the area coverage filter. This would make the availability distribution of time-series less irregular, therefor, make possible to create a model more prone to generalization both in space and time.

4. CONCLUDING REMARKS

This study has successfully tested the idea of using a DL approach for recognizing temporal signatures in time-series extracted from the high-res PlanetScope imagery. This open up a wide range of possibilities, one of them is identifying LCLU classes that need a more detailed information in order to be identified. Another possible application would be the ability to accurately identify change detection (e.g. deforestation). But there's no free meal! The great benefits of being able to deal with high-res RS data comes at the cost of making the very high volume of available information to be useful for creating models that are capable to generalize in both space and time. As observed, there's a great unbalance in the observations availability distribution between distinct instances of analysis, such that, inevitably, a good generalizable approach has to be able to consider gaps of information in order to be useful in many scenarios.

ACKNOWLEDGMENTS

This work, under the NexGenMap and MapBiomas initiatives (http://mapbiomas.org), was supported by the Gordon and Betty Moore Foundation (GBMF), The Nature Conservancy (TNC), WWF Brazil, the State of Goiás Research Foundation (FAPEG) and the Brazilian Council for the Scientific and Technological Development (CNPq).

5. REFERENCES

[1] Yan, J.; Ma, Y.; Wang, L.; Choo, K.K.R. and Jie, W., "A cloud-based remote sensing data production system", *Future Generation Computer Systems*, 86 (v.), 1154-1166 (pp.), 2018.

[2] A. Marinan, A. Nicholas and K. Cahoy, "Ad hoc CubeSat constellations: Secondary launch coverage and distribution", *2013 IEEE Aerospace Conference*, (pp.) 1-15., 2013.

[3] Houborg, R. and McCabe, M.F., "Cubesat enabled Spatio-Temporal Enhancement Method (CESTEM) utilizing Planet, Landsat and MODIS data." *Remote Sensing of Environment*, 209 (v.), 211-226 (pp.), 2018.

[4] LeCun Y., Bengio Y. and Hinton G., "Deep learning. ", *Nature*, 521 (v.), 436-444 (pp.), 2015.

[5] Kussul, N., Lavreniuk, M., Skakun, S., and Shelestov, "A. Deep learning classification of land cover and crop types using remote sensing data", *IEEE Geoscience and Remote Sensing Letters*, 14 (v.), 778-782 (pp.), 2017.