Sequential Recurrent Encoders for Land Cover Mapping in the Brazilian Amazon using MODIS Imagery and Auxiliary Datasets

Alejandro Coca-Castro[1,3] Marc Rußwurm[2] Louis Reymondin[3] Mark Mulligan[1]

[1]King's College London, Department of Geography [2]Technical University of Munich, Chair of Remote Sensing Technology [3]International Center for Tropical Agriculture (CIAT), Decision and Policy Analysis

{alejandro.coca_castro,mark.mulligan}@kcl.ac.uk, marc.russwurm@tum.de, l.reymondin@cgiar.org

Introduction

- **Land use and cover (LUC)** mapping approaches often require pre-processed cloud free observations.
- **End-to-end (E2E)** architectures based on **Recurrent Neural Networks (RNNs)** have shown promising signs for extracting relevant features from a **sequence of satellite image observations**.
- E2E models are able to learn relationships with **none or minimal region-specific expert knowledge**.
- These models and their variants are therefore **good candidates** to be tested for LUC classification across **large areas such as the Brazilian Amazon (BA)**.

Objectives

- To test an existing **sequential recurrent encoders** model based on **convolutional variants of RNNs** for the task of LUC classification across the Brazilian Amazon.
- To compare different **arrangements of input features** and their impact on the classifier performance.

Methodology

We oriented our implementation on a model design proposed by Rußwurm and Körner [2] (Figure 1). The essentials of our implementation are as follow.

- **A single-layer convolutional RNN bidirectionally encodes** a sequence of T satellite images to a fixed length representation (latent space).
- A convolution compresses the depth of the concatenated final outputs to **softmax-normalized activation** maps related to the number of target classes.
- Gradients from the cross entropy loss between prediction and ground truth are back-propagated and adjust the network weights using the **Adam optimizer**.
- Tune-able **hyper-parameters** are the **number of recurrent cells** r, the **sizes of the convolutional kernel** k_{rnn}, **the classification kernel** k_{class}, and **type of ConvRNN-cell** (*ConvGRU* [1] or *ConvLSTM* [3]).

We trained the architecture using a single layer of 64 cells of ConvGRUs, and k_{rnn} and k_{class} sizes of 3x3 pixels.

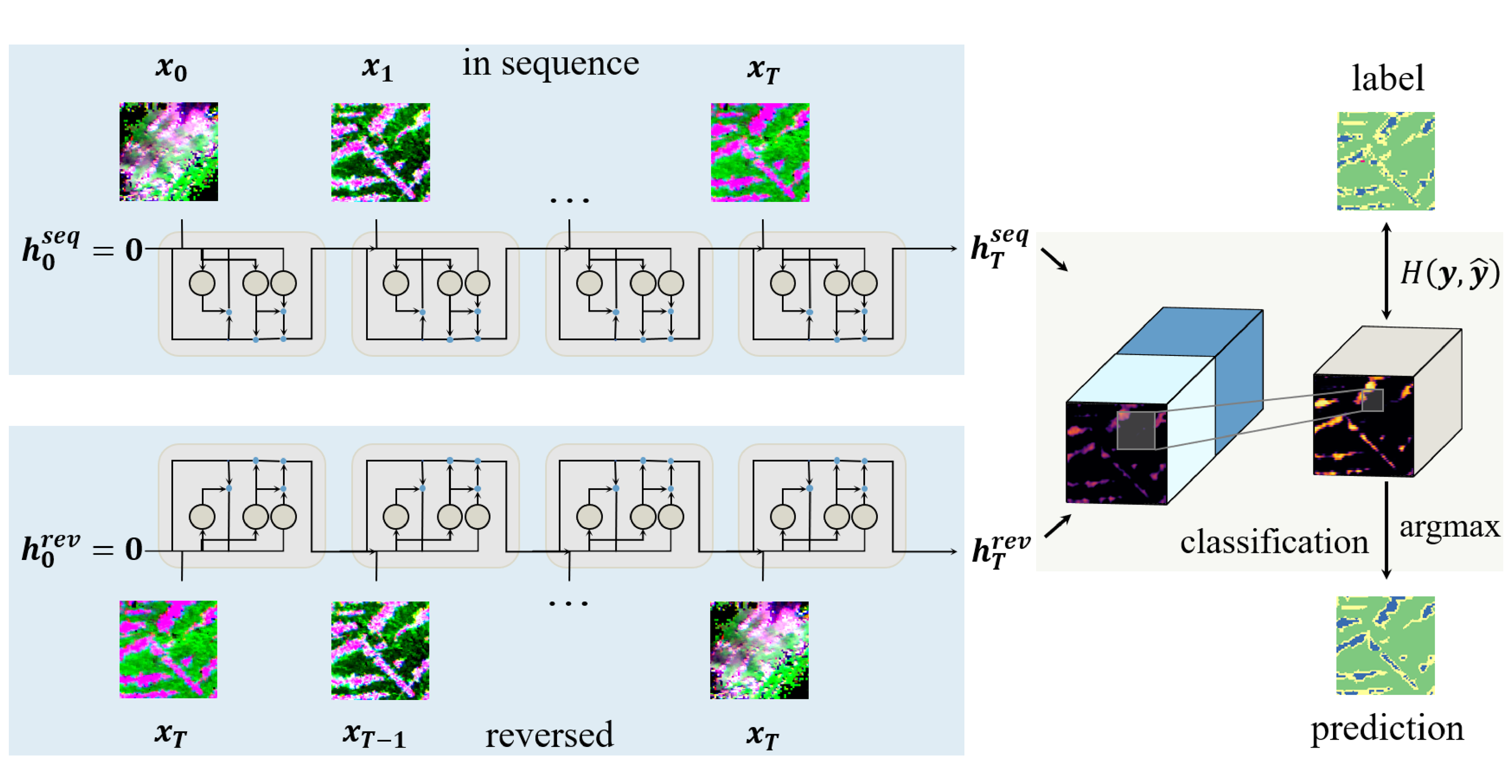

Figure 1: Adapted scheme of the architecture based on RNNs tested in this work. Original source: Rußwurm and Körner [2]

Study Area and Data

The Brazilian Amazon biome represents a large region covering approximately 4.1 million km^2 in South America. A **training set with images taken in 2009** were extracted from the *Google Earth Engine (GEE)* platform.

- Satellite data
  - MODIS Terra 8-day composite (MOD09Q1) Collection 6 250-m (**NIR, MIR**)
  - MODIS Terra 8-day composite (MOD09A1) Collection 6 500-m (**blue, green, SWIR1, SWIR2, SWIR3**)
- Auxiliary data
  - CGIAR SRTM Data V4 90-m (**elevation**)
  - WordClim Bio Variables V1 1-km (**annual avg temperature and precipitation**)
- Reference data (labels)
  - **MODIS land cover product** (MCD12Q1) Collection 6 500-m (**17 classes**)

For data partition, blocks of 384x384 pixels were created across the area of interest. The block size matches multiples of **24x24 pixels** which was the **patch size** used for model training, testing and evaluation under a ratio of 4:1:1 (59M, 15M, 15M pixels respectively).

Results

In overall, the **auxiliary data and spectral indices** can **contribute positively and negatively, respectively** to the LUC classification at the study area (Table 1).

Table 1: Weighted average values of pixel-wise metrics, including overall accuracy (OA), of the trained convolutional GRU sequential encoder networks after training **100 epochs**.

Data	Short name	OA	Kappa	Precision	Recall	F1-Score
Spectral Reflectance	SR + TI	0.98	**0.87**	**0.95**	**0.95**	**0.95**
Spectral Reflectance + Spectral indices*	SR + TI + SI	0.95	0.73	0.90	0.89	0.89
Spectral Reflectance + Auxiliary Data	SR + TI + AUX	**0.99**	**0.87**	**0.95**	**0.95**	**0.95**
All (SR + TI + SI + AUX)	ALL	0.96	0.73	0.9	0.89	0.89

* The indices extracted include the Two band Enhanced Vegetation Index (EVI2), Normalized Difference Wetness Index (NDWI), Normalized Difference Infrared Index with SWIR2 (NDII1) and Normalized Difference Infrared Index with SWIR3 (NDII2).

The kappa metric normalizes for imbalanced classes and is thus used to evaluate the classification (Figure 2).

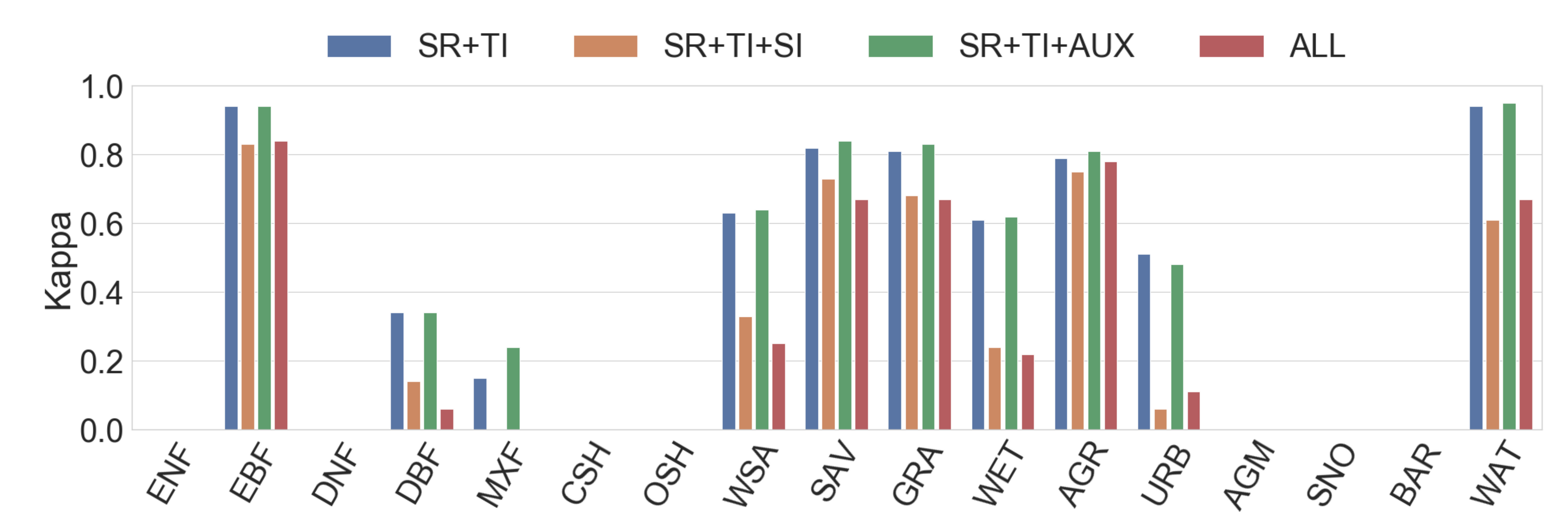

Figure 2: Kappa values for LULC classification using trained convolutional GRU sequential encoder networks after training 100 epochs on data with different input features.Abbreviations of the classes include Evergreen Needleleaf Forests (ENF), Evergreen Broadleaf Forests (EBF), Deciduous Needleleaf Forests (DNF), Deciduous Broadleaf Forests (DBF), Mixed Forests (MXF), Closed Shrublands (CSH), Open Shrublands (OSH), Woody Savannas (WSA), Savannas (SAV), Grasslands (GRA), Permanent Wetlands (WET), Croplands (AGR), Urban and Built-up Lands (URB), Cropland/Natural Vegetation Mosaics (AGM), Permanent Snow and Ice (SNO), and Barren (BAR).

The spatial patterns of the predictions were examined over a representative evaluation block of 384x384 pixels (Figure 3).

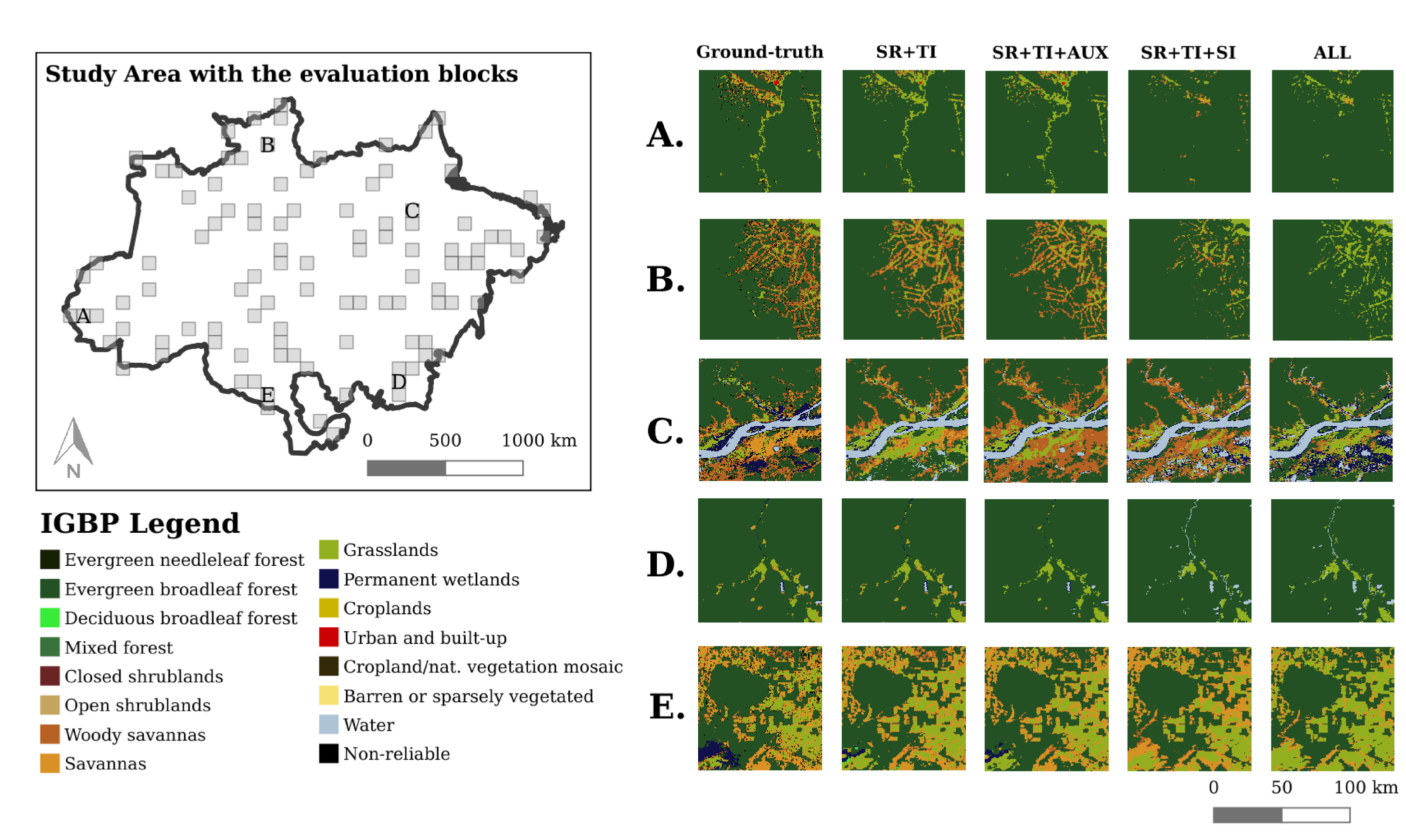

Figure 3: Qualitative results of the inference for 2009 by the trained convolutional GRU sequential encoder networks using different input features. The map insets correspond to five evaluation blocks distributed across the study area.

Conclusions

- A state-of-art **model based on RNNs was successfully assessed for large-area classification** across BA using MODIS archives.
- The architecture achieved satisfactory results with or without using auxiliary data.
- Spectral indices affected the performance of the classifier.
- The straightforward nature and properties (ie., **minimal preprocessing and no additional cloud filtering**) of the network remain **relevant to continuous LUC monitoring** over the study area.
- **Future work** includes to **investigate the cloud filtering** feature of the model and **performance using other reference datasets** across BA.

References

[1] N. Ballas, L. Yao, C. Pal, and A. Courville. Delving deeper into convolutional networks for learning video representations. *arXiv preprint arXiv:1511.06432*, 2015.

[2] M. Rußwurm and M. Körner. Multi-temporal land cover classification with sequential recurrent encoders. *ISPRS International Journal of Geo-Information*, 7(4), 2018.

[3] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *Advances in Neural Information Processing Systems 28*, 2015.

– Acknowledgements: authors acknowledge the International Center for Tropical Agriculture (CIAT) for granting the access to the resources used for processing data and models linked with this work –