Deep learning methodology for predicting socioeconomic indicators in Vale do Ribeira using satellite imagery

Isabella de Melo Sousa[1], M. Jeaneth Machicao Justo[2], and Pedro Corrêa[2]

[1]Escola Politécnica da Universidade de São Paulo
[2]Universidade de São Paulo

November 21, 2022

Abstract

Key measures of socioeconomic indicators are essential for making informed policy decisions, but due to the high costs and operational difficulties of traditional data collection efforts, obtaining reliable socioeconomic data remains a challenge, particularly in developing countries. This work presents a deep learning methodology to estimate socioeconomic indicators using satellite imagery. The neural network model developed was trained at the Brazilian region of Vale do Ribeira with the goal of analyzing the socioeconomic indicator of income. The preliminary results showed that models using nightlight (NL) or multispectral daytime (MS) imagery performed better than models trained only on RGB bands and that models trained exclusively on NL or MS imagery performed similar to one another and nearly as well as the combined model MS+NL. Finally, the model yielded a low performance (R2 = 0.289), but it is still promising once the dataset employed was considerably smaller than the one used in the original study that attempted to replicate.

Deep learning methodology for predicting socioeconomic indicators in Vale do Ribeira using satellite imagery

SOUSA, I.[1]; MACHICAO, J.[1]; CORREA, P.[1]

[1] Computer Engineering Department - Escola Politécnica da Universidade de São Paulo

Email: {isabellamsousa, machicao, pedro.correa}@usp.br

INTRODUCTION

Key measures of socioeconomic indicators are essential for making informed policy decisions, but due to the high costs and operational difficulties of traditional data collection efforts, obtaining reliable socioeconomic data remains a challenge, especially in developing countries. In the Brazilian scenario, the main census organization is called the Brazilian Institute of Geography and Statistics (IBGE), and data collection typically occurs every ten years. Due to the data gap created by this frequency, the results of the collection may not reflect Brazilian reality in real-time, harming the interpretability of the results. In order to fill such a gap, this research aims to develop a low-cost and scalable deep learning method for estimating socioeconomic indices using satellite imagery.

MATERIALS AND METHODS

The article of reference adopted for the development of this work was the paper of Yeh et al. (2020) [1], that combines both multispectral daytime imagery (MS) and nightlights imagery (NL) in a deep learning model trained end-to-end to estimate socioeconomic criteria in several regions of Africa. The Vale do Ribeira region was used as a case study to replicate the paper's techniques in the Brazilian context.

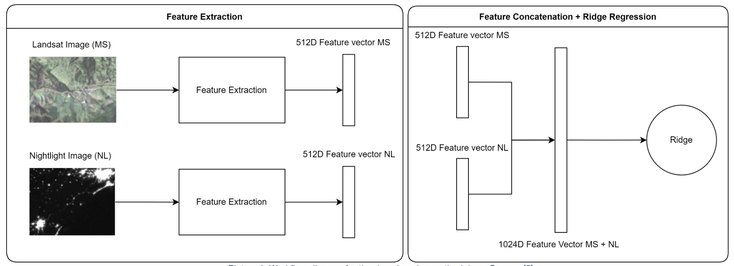
The workflow of the project was divided into four steps [2]:

- Step 1: Satellite imagery and IBGE census data acquisition and aggregation.

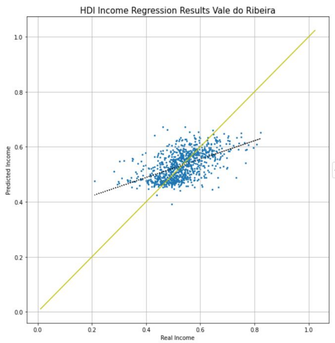
Vale do Ribeira's census sectors were chosen as the model's level of granularity. The socioeconomic variable (income) was provided by the IBGE from the census of 2010, and it was adjusted to the proper granularity by utilizing the methods of Abreu et al [3]. The satellite imagery was collected using the Google Earth Engine API.

- Step 2: Feature extraction

Two pretrained Resnet-18 networks models were modified to adapt multi-band satellite imagery and used to extract the feature vectors of the images. The loss function used was mean squared error (MSE). The training followed a 5-fold cross-validation.



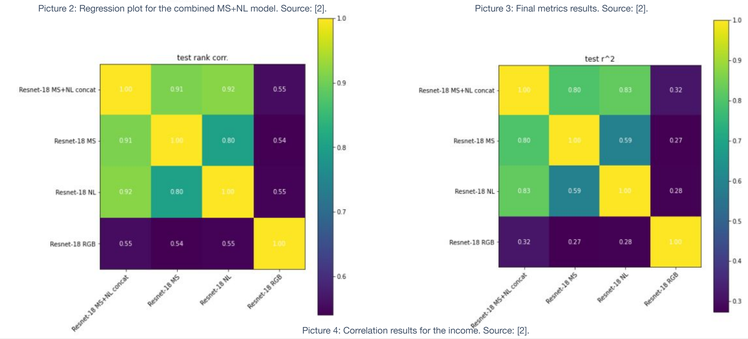
Picture 1: Workflow diagram for the deep learning methodology. Source: [2].



Picture 2: Regression plot for the combined MS+NL model. Source: [2].

	r2	R2	mse	rank
Resnet-18 MS+NL concat	0.294940	0.289370	0.004663	0.567761
Resnet-18 MS	0.285869	0.285763	0.004686	0.556179
Resnet-18 NL	0.230912	0.223876	0.005092	0.521667
Resnet-18 RGB	0.110392	0.109057	0.005846	0.298294

Picture 3: Final metrics results. Source: [2].



Picture 4: Correlation results for the income. Source: [2].

- Step 3: Ridge regression

The feature vectors were fed to the fully connected layer of the Resnet-18. The loss function used was MSE. The training followed a leave-one-group-out cross-validation.

- Step 4: Results Analysis

Performance metrics such as the coefficient of determination (R^2) were calculated, and graphs were drawn to better analyze the results.

RESULTS

- Comparing to the results obtained by Yeh et al (R^2 = 0.70), the model performs poorly (R^2 = 0.289). Nevertheless, it is still promising once the dataset utilized was significantly smaller than the one used in the original study (4.4%);
- In contrast to the other models, the model trained only on the RGB bands performed badly, which supports the relevance of employing multispectral daytime and nighttime imagery when predicting socioeconomic variables;
- The models trained exclusively on NL or MS imagery performed similar to one another and nearly as well as the combined model MS+NL, indicating that the information contained in these two inputs is similar.

CONCLUSIONS

The idea of using deep learning to estimating socioeconomic variables is very promising, and although the model did not perform well, this study is a step towards understanding how convolutional neural networks and satellite imagery can be used to this end.

References:

[1] YEH, C. et al. Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. Nat Commun 11, 2583 (2020). https://doi.org/10.1038/s41467-020-16185-w

[2] SOUSA, I. A deep learning approach to predict socioeconomic indicators in vale do ribeira from satellite imagery. https://zenodo.org/record/6366429#.YxZC13bMK3B

[3] Abreu, Marcos & Oliveira, Julio & Andrade, Viviane & Meira, Anderson.(2011). Methodological proposal for spatial calculation and analysis of the intra-urban HDI of Viçosa, Brazil. Revista Brasileira de Estudos de População. 28. 169-186..

Acknowledgements:

I would like to thank Prof. Dr. Pedro Corrêa, my teacher and supervisor.

Ph.D Jeaneth Machicao, for whom I express my sincere gratitude for her availability and advice.

And to the PARSEC group, in which this work is part of.

Author affiliations: Isabella de Melo Sousa (University of São Paulo, BR); Jeaneth Machicao (University of São Paulo, BR); Pedro Luiz Pizzigatti Corrêa (University of São Paulo, BR).