Detecting Urban Deforestation: A Semantic Segmentation Approach

Ilan Grynszpan[1], João Pedro Jesus de Abreu Martinez[1], Pedro Ivo Mioni Camarinha[2]

[1]Pontifical Catholic University of Rio de Janeiro - PUC-Rio
[2]National Center for Monitoring and Early Warning of Natural Disasters – CEMADEN

Keywords: Remote Sensing, Deep Learning, Urban Deforestation, Urban Heat Island, Rio de Janeiro

Abstract

The increase in built-up areas within metropolitan perimeters has far-reaching negative effects. The increase in buildings, suburbs, and roads leads to reductions in native environment areas, contributing to the phenomenon of urban heat-islands (UHIs). According to the World Bank (World Bank Group, 2023), the temperature of south-eastern Asian cities is on average 1.6 to 2.0 degrees warmer than their surroundings, with some cities reaching 5.9 degrees warmer than the surrounding areas. Extreme urban heat has negative consequences including reducing city Gross Domestic Product (GDP) and increasing the demand for electric energy. Furthermore, it can lead to the accumulation of greenhouse gases in the local area and can cause heat-related health issues, which can even result in death (Environmental Protection Agency, 2008). Measuring the decrease of natural areas within cities is a complicated task which involves government monitoring and accurate mapping. This work aims to accurately detect the reduction in natural areas within the city of Rio de Janeiro using semantic segmentation machine learning techniques. The models used were able to detect vegetation and urban areas with over 90% accuracy.

1. Introduction

1.1 Context

The share of the world´s population living in cities is increasing every year. According to the United Nations (UN, 2025), 55% of the planet´s population currently live in urban areas and this is predicted to increase to 68% by 2050. The significant population increase in urban areas requires more and more urban infrastructure developments thus leading to the construction of new houses, buildings, suburbs, and roads.

The construction of new suburbs and the increase in size of existing ones have good economic outcomes through the generation of jobs, the creation of new commercial opportunities, and by increasing economic growth. However, unorderly urban sprawl can have severe economic and environmental consequences (Balaban, 2012). For example, Wolan and Schick (1965) cite the rapid soil erosion in built-up areas as one such consequence. Another negative consequence is the formation of Urban Heat Islands (UHIs) which result in temperatures in urban areas being significantly hotter than their rural surroundings and contribute to the concentration of greenhouse gases in these locations (Memon, 2007).

The phenomena of UHIs occur when the temperature in an urban area is much higher than the temperature in its surroundings. This is a result of heat generated by solar radiation as well as by the heat generated by power-consuming devices such as air conditioners. These sources create heat which cannot be dissipated easily in urban areas due to insufficient levels of vegetation, tree cover and various other factors. According to Phelan et al (2015), UHIs have far-reaching implications such as increasing energy consumption due to the need for air-conditioning, decreasing the quality of water due to warming of water sources, and endangering native species.

Another pervasive effect of the concentration of buildings in urban areas is the accumulation of greenhouse gases in these areas which worsens air quality and may cause long-term diseases among the local population. Furthermore, the occurrence of UHIs can increase the effects of heat waves, leading to hospitalization and death, especially among older individuals. Additionally, the accumulation of greenhouse gases has been linked to an increase in the rate of respiratory diseases, which is another negative health consequence (Phelan et al, 2015). According to Phelan et al (2015), possible mitigation strategies for the UHI effect include changing asphalt material to more reflective options, and increasing the amount of vegetation within urban areas through the establishment of parks and green roofs.

Furthermore, some institutions have studied the mapping of urban areas and have generated datasets, maps and studies which can be useful in further urban expansion planning. One example is the institute Data Rio, which is linked to the City Government of Rio de Janeiro. Another example is the Institute MapBiomas (MapBiomas, 2025) which has mapped the entire region of Brazil and has captured a high diversity of biomes and urban areas.

A great motivator for this work was the mapping undertaken by Data Rio which examined the vegetation coverage and land use classifications in 2010, 2016, 2018 in Rio de Janeiro. Although Data Rio´s methodology for the creation of their maps is not publicly available, they have provided maps for the years aforementioned to the public on their website. After the researchers examined Data Rio´s publicly available map archives, it became apparent that Data Rio had not created any further maps about land use and land coverage after the year 2018 in Rio de Janeiro. This may be because generating updated maps containing these analyses each year for an area with approximately 1,221 km^2 is highly time consuming, labour intensive and requires efforts from highly qualified staff with geospatial and cartographic knowledge.

Although it is not known which methodology Data Rio used to generate their maps, elsewhere in the literature deep learning techniques for semantic segmentation have been used to

analyze and then create maps detailing Land Use Land Cover (LULC) (Zhang et al., 2019). Furthermore, some studies have generated maps of LULC using images from the Sentinel-2 satellite (Campos-Taberner et al., 2020). Due to these factors, we decided to combine the current Data Rio datasets with optical satellite images from the Sentinel-2 satellite which are available through Google Earth Engine (GEE). As the data from Data Rio and the satellite images are both publicly available, we could build a deep learning semantic segmentation pipeline to complete the missing yearly maps of land use and vegetation coverage for the municipality of Rio de Janeiro from 2019 to 2024.

2. Motivation

To further explore the relationship between the formation of UHIs and built-up urban areas in Rio de Janeiro, we examined data from a study conducted previously by Data Rio (Prefeitura da Cidade do Rio de Janeiro, 2016b). In this study, data was collected on various dates from a thermal infrared sensor (TIRS) via the satellite Landsat 8 to estimate surface temperatures at a resolution of one-hundred meters between 2013 and 2016. The results of this study reinforced the evidence that more densely urbanized places tend to form UHIs more than areas with greater vegetation cover. Therefore, with the predicted increases in urban populations over the next few decades it is important for researchers and government authorities to monitor and predict the sprawl of urban areas and building density in order to avoid the negative effects of UHIs as previously mentioned.

Other sources of remote sensing maps including both urban areas and areas of preserved vegetation in Rio de Janeiro are the datasets provided by Data Rio, known as *Cobertura Vegetal e Uso da Terra*. This data is available for the years 2010, 2016 and 2018 ((Prefeitura da Cidade do Rio de Janeiro, 2018), (Prefeitura da Cidade do Rio de Janeiro, 2016a)). These datasets present multiple land use classes related to anthropogenic activities and also include different classes of natural vegetation within the metropolitan area. However, after 2018 no further maps were produced by Data Rio which until this current work meant that monitoring the changes in vegetation cover and the planning of urban zones to avoid UHIs were not being undertaken.

Additionally, Institute MapBiomas has also conducted studies to classify urban and naturally preserved areas (MapBiomas, 2025). With an extensive spatio-temporal series at 30 meters resolution from Landsat imagery, the institute provides consistent LULC maps which can identify trends and detect changes in the LULC. Besides this, MapBiomas also publishes their methodological approaches, detailing their use of deep learning models to improve the classification accuracy of complex classes. Therefore, we decided to use the MapBiomas results as a baseline to discuss the results found in our current study.

MapBiomas offers highly detailed maps of the urban development of Rio de Janeiro, however, it has been noted that Data Rio has been able to better capture different features (i.e. exposed soil) that are relevant to the analysis of urban development. The inclusion of these features and the possibility of a different methodology used by Data Rio provides a novel way to map urban development. However, as Data Rio has not released any updated maps since 2018, this justifies the need for the completion of subsequent maps which we aim to do in this study.

The first point we considered before performing a semantic segmentation trained on the Data Rio spatial datasets was that some studies have suggested that MapBiomas fails to adequately capture certain features of the landscape within their maps such as exposed soil (Louzada, 2025). In addition to this, the maps from MapBiomas are developed using similar technologies such as U-Net (Ronneberger et al., 2015), however, these maps capture the entire landmass of the country of Brazil. Therefore, it is assumed that MapBiomas has trained their models with a much broader set of samples and classes to enable the mapping of several different environments, not only the specific ones for the urban environment of Rio de Janeiro.

We explored MapBiomas´s technical documentation for several collections (MapBiomas Project, 2024, 2023, and 2022) which showed that the company uses several different sources to train their models, including previous collections of their own service. However, it is important to note that data from the City Government of Rio de Janeiro is not available as one of these sources. Generating the maps of LULC according to the Data Rio methodology for the years after 2018 is also important in order to compare our results with the maps created by MapBiomas.

Furthermore, MapBiomas has different definitions of urban areas than the ones used in our study. The proper definition of which classes from Data Rio were used as "urban", "vegetation", and "other" is detailed in section 3, but for the purpose of comparison, MapBiomas (MapBiomas, 2025) has the following subclasses under "non-vegetation area": ***Beach, dunes, and sand, Urban areas, Mining areas, Other non-vegetation areas***. In addition to this, if the settings in the MapBiomas maps are changed to show "anthropism" areas, the subclasses are: ***Pasture, Mosaic of land uses, Urban areas, Mining areas, Aquaculture***.

In this study the definition of the "urban" class takes into consideration the following original subclasses present in the Data Rio datasets: ***Urban areas, Mining areas, Exposed soil***. It is important to note that the "exposed soil" subclass does not appear on MapBiomas maps. However, several of the classes specified by MapBiomas are not categorized as "urban" areas in this study therefore leading to less areas classified as urban. This classification was the most appropriate because of the reported purpose of this research: to help prevent the UHI effect, which is caused by urban development and not by farmland and other kinds of anthropogenic activity involving the natural environment.

Some studies suggest that "mining areas" could impact on UHI formation (Zhenqi, 2017). Based on this, this original class was further aggregated into the "urban" class. Moreover, the inclusion of "exposed soil" within the urban sprawl mapping was debated while processing the dataset´s classes. As better described in future sections of this paper, there were no considerable occurrences of "exposed soil" in the samples obtained from Data Rio. Therefore, it was decided that the "exposed soil" class could be used as a proxy for further urban development either for mining or for construction as both of these factors can contribute to the formation of UHIs.

3. Methodology

3.1 Datasets

To train our semantic segmentation deep learning model, two datasets were required: (1) the Data Rio datasets called *Cober-*

tura Vegetal e Uso da Terra and (2) the satellite imagery from the Sentinel-2 MultiSpectral Instrument (MSI), collected via Google Earth Engine (GEE) through the Harmonized Sentinel-2 Level-1C (L1C) collection (Earth Engine Data Catalog, 2025). This decision followed an evaluation of the available satellite data options, in which Sentinel-2 was identified as being most suitable as it offered publicly accessible imagery at a high spatial resolution.

The L1C products provide top-of-atmosphere reflectance values, radiometrically corrected and orthorectified using a digital elevation model in UTM/WGS84 projection. However, although this guarantees geometric consistency across time, it does not correct for atmospheric effects such as water vapor, aerosols, and haze, which introduce noise into the reflectance values. Conversely, the Sentinel-2 Level-2A (L2A) provides surface reflectance, or bottom-of-atmosphere reflectance derived from L1C through atmospheric correction (Copernicus Data Space Ecosystem, 2025).

Despite the acknowledged advantages of L2A, the present study employed L1C imagery. The availability of L1C imagery in the GEE collection aligns better with the Data Rio reference shapefiles, offering coverage from June 2015, whereas the L2A collection in GEE begins only in March 2017. In this sense, while the use of L1C already excludes the 2010 Data Rio dataset from consideration, relying solely on L2A would further exclude the 2016 dataset, thereby reducing the training samples available for this study.

It is also worth noting that L2A products can be generated from L1C products using the Sentinel-2 Toolbox with processors like Sen2Cor for atmospheric correction (European Space Agency, 2025). However, this would require downloading and preprocessing large volumes of data outside GEE which reduces reproducibility. For this study we prioritized reproducibility with publicly accessible datasets and consistent temporal coverage, accepting the trade-off of retaining some atmospheric influence in top-of-atmosphere reflectance values.

In order to construct the satellite images, three bands were selected from the Sentinel-2 MSI: B4, B3 and B2. The equivalent wavelengths are described below:

- B4: 664.6 nm (S2A) / 664.9 nm (S2B) - Red
- B3: 559.8 nm (S2A) / 558.9 nm (S2B) - Green
- B2: 492.7 nm (S2A) / 492.3 nm (S2B) - Blue

There is a slight difference between the twin satellites Sentinel-2A (S2A) and Sentinel-2B (S2B) (European Space Agency / Copernicus SentiWiki, 2025)

Bands B4, B3, and B2 fall within the visible spectrum and correspond directly to the wavelengths perceived by the human eye as red (R), green (G) and blue (B) light respectively. By assigning each spectral band to its matching RGB channel, we created true-color composites that emulate natural human vision, facilitating the interpretation of urban and vegetated areas.

The imagery was filtered by cloud cover, retaining only scenes with less than 10% cloudiness, which were then aggregated using their temporal mean. Additionally, each RGB channel was independently normalized using min–max scaling, adjusted according to its own value range, and followed by a gamma correction to balance the overall intensity across the composite.

The Data Rio datasets initially provided us with shapefiles containing geometries of the LULC classes. Since this paper proposes to identify the urban development with a focus on mitigating the establishment of future UHIs, the various classes were grouped into super-classes **{urban, vegetation, others}** and the geometries were reclassified. More details about which super-class the original classes were grouped into are presented in Table 1.

As a final step, a grid was built at 10 meters resolution, comprising the entire area of Rio de Janeiro municipality to process the super-class geometries and ensure they were compatible with the satellite-processed images. To achieve this, we employed rastering methods that served to georeference the geometries from the 2016 and 2018 Data Rio datasets in this grid.

This process resulted in two separate reference maps for each year. A numerical label for each pixel was assigned: (1) for **vegetation** cover; (2) for **other** classes, such as beaches and mountains; (3) for **urban areas**. Figure 1 shows each of these classes in the processed grid as indicated by the different colors.

Figure 1. The processed reference map for Data Rio 2018 dataset at 10 meters resolution showing the super-classes: ”vegetation” (purple), ”other” (orange), and ”urban” (yellow).

3.2 Model

The model chosen to work with semantic segmentation of the LULC was the U-net. It presents a good track record with this type of task due to its skip-connections which helps the model to recover previous layers' feature maps from encoder layers. It integrates them to feature upsampled feature maps of the decoder, mitigating details lost in the encoding of the inputed images through dimensionality reduction of the sequential convolutional layers.

The encoder structure used was a ResNet (He et al., 2016). This architecture is designed to mainly avoid the vanishing gradients problem. This is done internally by creating a residual layer for the gradient, and thus an identity mapping structure allowing for not only a nonlinear representation of the identity function but also a better network learning.

3.3 Experimental Protocol

For this paper, we decided to proceed with a dual training protocol before predicting the missing yearly maps for land use and vegetation cover from Data Rio.

At first the satellite images from 2016 were used to train the model and the mapping of classes (urban, vegetation, or other areas) from Data Rio was presented to the model as references.

Urban	Vegetation	Others
Urban (construction-filled) areas	Dense Ombrophilous Forest	Rocks (Rocky outcrop)
Mining areas	Pioneer formations with marine influence	Agricultural land
Exposed soil	Pioneer formation with fluvial-marine influence	Continental bodies of water
	Pioneering Formation with fluvial-lacustrine influence	Coastal bodies of water
	Re-forestation	Beach
	Non-forest tree vegetation	
	Grassland vegetation	

Table 1. Super-classes derived from the aggregation of original land use and vegetation cover classes in the Data Rio datasets (2016–2018), based on the reference geometries.

After this training was finished, the results for the 2018 satellite images were compared to the 2018 class map to evaluate the performance of the model. After this procedure was finished, the images from 2018 were used to train the model which was then tested a second time using the 2016 images.

The major goal for this dual-training was to assess predictive metrics for semantic segmentation to determine if the model could generalize different inferences across time periods. Based on these results, we trained the model for a third and final time. We used both the 2016 and 2018 satellite images to train the model and predict labels for out-of-sample imagery from 2019 until 2024.

In addition, the models were trained using the Adam optimization algorithm, starting with a learning rate of 1e-3 and a weight decay of 1e-5. To adjust the learning rate during training, we adopted a cosine annealing schedule with a lower bound of 5e-5. The training process was constrained to 200 epochs with early stopping applied based on validation loss using a patience threshold of 10 epochs to mitigate overfitting.

The loss function used was the *cross entropy*, a common function which is used in classification applications (Jadon, 2020). Although the authors acknowledge that functions like the *focal loss* are better for some kind of class imbalancing, the main idea was to identify the class of urban areas, which is correctly balanced with natural environments. The loss function is the following, where $p(i)$ is the probability assigned for the true class of the pixel:

$$CE = -\frac{1}{N} \times \sum_{i} log(p_i) \tag{1}$$

To build our samples to train the model, we first extracted patches of 128x128 dimensions following the condition of having at least one pixel belonging to the studied area raster, also using a 60% overlap to balance the spatial coverage. Given that, we obtained a total of 5401 patches in which 70% were assigned for training and 30% for validation. This amount doubled during the final training using both the 2016 and 2018 datasets.

Each sample contained an input-target tensor pair with dimensions of (128, 128, 3) and (128, 128, 1), respectively. This was due to the nature of the input, which corresponded to the processed RGB bands of Sentinel-2 satellite images, cropped at the spatial reference provided by a patch. Conversely, the target tensors were single-channeled and had only the integer class labels provided by the reference grid maps of the processed Data Rio datasets.

3.4 Further considerations about class imbalance

Although it is not the focus of this paper, there may be issues with inter-class imbalancing, and these could be addressed in a future version. This paper did not intend to further differentiate urban expansion due to mining, urban sprawl, or exposed soil as individual components. However, when assessed, the proportions in relation to the total landmass of Rio de Janeiro of these components were dominated by ”Urban areas” (46.59%). ”Mining areas” (0.33%) and ”Exposed soil” (0.64%) contributed a much smaller percentage. Further investigations will be necessary to pinpoint the proportional contributions of each of these components in relation to the formation of UHIs. This would therefore justify a further division of the superclasses.

The superclass ”others” was under-represented in the sample. Despite this under-representation, cross-entropy loss function was used as the role of this paper was to capture the more prevalent ”vegetation” and ”urban” classes. It is possible that the misclassification of the class ”others” could incorrectly indicate an increase in urban areas. However, this effect was not verified.

4. Experimental Analysis

Initially, we focused our analysis on well-established semantic segmentation metrics to assess the performance of the conducted dual-training: (1) we trained the U-net model using satellite images from 2016 and tested it with Data Rio processed reference LCLU maps; (2) we later trained and tested using the 2018 dataset for training and the 2016 dataset for testing.

As shown in Tables 2 and 3, the first step of the dual-training procedure consistently outperformed the second step in nearly all evaluation metrics. Notably, the first step achieved higher global Intersection Over Union (IoU) (0.756 vs. 0.649) and overall pixel accuracy (0.915 vs. 0.878) indicating better spatial alignment and overall classification performance. Class-wise, both vegetation and urban categories maintained strong predictive capacity in both training steps with the first step showing slightly superior F1-scores (0.919 and 0.942, respectively) compared to the second step (0.881 and 0.920).

The most significant drop in performance occurred in the “other” class with a decrease from 0.528 to 0.306 on its IoU and from 0.691 to 0.469 on its F1-score. We argue that the main reason for this was a general imbalance between the class ”other” and the classes ”vegetation” and ”urban”, resulting in a lower generalization for the former class. In addition, we raise the hypothesis that the second step of dual-training, trained using the

Dual-Training - 1st Step - Training: 2016, Testing: 2018					
	IoU	Precision	Recall	F1-Score	Pixel Accuracy
global	0,756	-	-	-	0,915
vegetation	0,850	0,889	0,950	0,919	-
other	0,528	0,788	0,615	0,691	-
urban	0,890	0,960	0,924	0,942	-

Table 2. Inference metrics assessed over 2018 after training over 2016 satellite imagery.

Dual-Training - 2st Step - Training: 2016, Testing: 2018					
	IoU	Precision	Recall	F1-Score	Pixel Accuracy
global	0,649	-	-	-	0,878
vegetation	0,788	0,893	0,870	0,881	-
other	0,306	0,635	0,372	0,469	-
urban	0,852	0,882	0,962	0,920	-

Table 3. Inference metrics assessed over 2016 after training over 2018 satellite imagery.

the satellite images from 2018, could have suffered from a non-optimal gamma correction thus making it more difficult for the model to distinguish ”other” pixels from ”vegetation” ones, e.g. rock formations near dense vegetation areas of a darker tone.

Nevertheless, both steps demonstrated solid predictive performance through mapping the RGB satellite imagery into land use and land cover classification. Figure 2 illustrates this behavior by presenting a mosaic with different spatial regions from Rio de Janeiro, where the pixel-wise label predictions closely matched the reference maps, with only minor misclassifications.

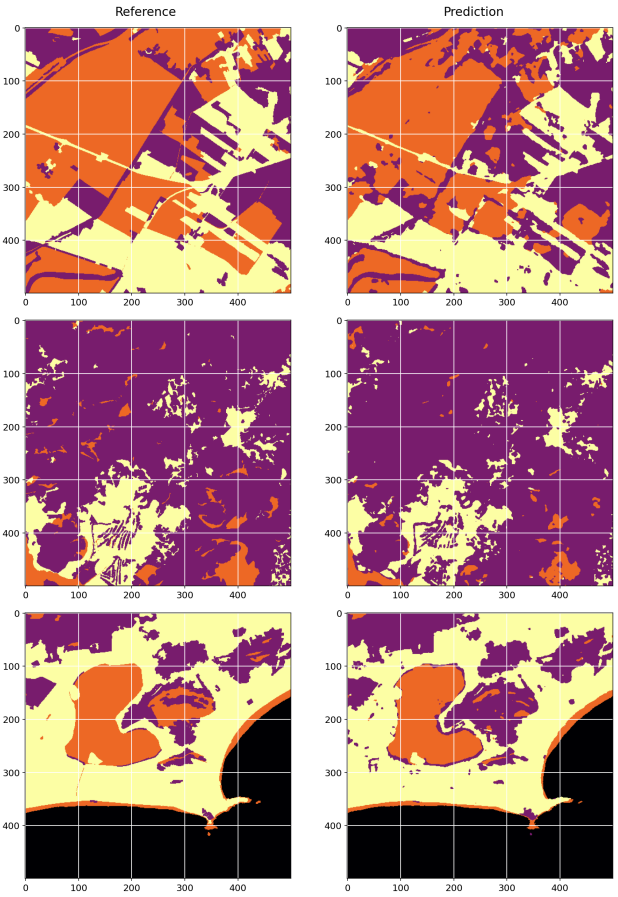

Figure 2. Mosaic comparing different scenes taken from both reference (left) and in-sample prediction maps (right), in 2018. Classes: vegetation (purple), urban (yellow), other (orange).

Given these results, we proceeded with our initial idea to combine the 2016 and 2018 satellite imagery and train a final model to predict the missing yearly maps from Data Rio, from 2019 to 2024. By counting the pixel labels in these predicted maps we produced a temporal evaluation of either increasing or decreasing trends of the total area for each super-class, as illustrated in Figure 3.

At first, we included the total areas retrieved from the reference maps of 2016 and 2018 (shown in solid lines) to capture the trends at these specific time instants. Accordingly, the chart presents increasing trends for ”urban” and ”other” areas at the expense of deforestation processes, reflected in the decreasing trend for ”vegetation” area. On the other hand, the total areas of the predicted maps (shown in dashed lines) exhibit an unexpected behavior, with ”vegetation” increasing while ”urban” and ”other” decrease. The only exceptions for these predicted trends occurred in 2021 and 2024.

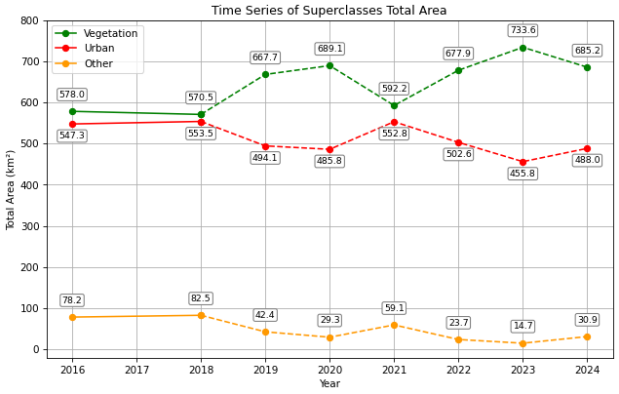

Figure 3. Time series of super-classes total area (km^2), including reference (solid lines, 2016 and 2018) and predicted (dashed lines, 2019-2024) values.

In Figure 4 we observe that the ”anthropism” class of MapBiomas presents a temporal evolution comparable to our model´s evolution for the super-class ”urban” during the same predicted years. It is also worth noting that the upwards and downwards trends observed in the MapBiomas data were also captured by our model albeit with a different scale. The upwards trend in 2021 was noticeable in both time series but at a much reduced scale in the MapBiomas trends.

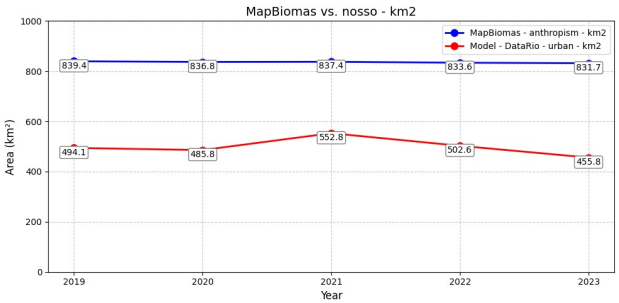

Figure 4. Time series comparison of total predicted area for ”urban” and ”anthropism” classes, from Data Rio and MapBiomas, respectively.

Conversely, our results showed an increase in vegetation between 2018 and 2021 and then between 2021 and 2024 that was not observed in MapBiomas. We further investigated the predicted

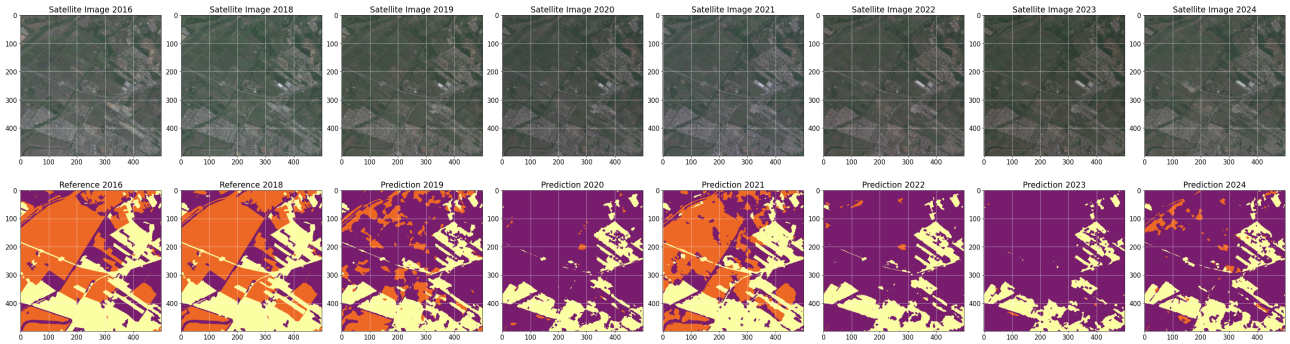

Figure 5. Mosaic with temporal evolution of a satellite scene and the corresponding references (2016 and 2018) and predicted maps (from 2019 to 2024), including all LULC super-classes—"vegetation" (purple), "urban" (yellow) and "other" (orange).

maps visually to better understand the results. Due to missing references for the predicted years, we based the following reasoning on the satellite imagery. Specifically, we considered observable urbanized areas and agricultural lands (representing the "other" super-class; see Table 1) in the 2016 and 2018 scenes, identified by their respective references. In case no significant changes occurred in the true-color visual patterns of these areas in the subsequent scenes, we would expect their LULC mapping to remain consistent with the earlier references.

In this context, another mosaic was built in Figure 5, showing the temporal evolution of a satellite scene alongside its references and predicted maps. For the 2016 and 2018 references, all super-classes were present, with emphasis on "urban" (yellow) and "other" (orange) areas. Observing the temporal evolution of the satellite images, the true-color visual patterns remained largely unchanged throughout the years. Conversely, the predicted maps showed the expansion of "vegetation" (purple) into areas previously mapped as "other". This misclassification applied to areas that resembled well-established agricultural lands. A similar issue occured for urbanized patterns: in the 2020 and 2023 predictions a highway and adjacent urban site that were visible in all satellite images within the vertical (200–300) and horizontal (0–300) ranges respectively were misclassified as "vegetation".

This explains the increasing trend of "vegetation" in Figure 3, which was inflated by misclassifications occurring mainly in the predicted periods of 2019-2020 and 2022-2023. However, when the predicted mappings in 2021 (Figure 5) were analyzed, there seemed to be less misclassification. The model more consistently classified visual patterns as "urban" and "other" for areas that had already belonged to those super-classes in earlier references; this remained nearly unchanged in the satellite images. Although the chart in Figure 3 showed a sudden decrease in "vegetation" in 2021, this new value represented a more reliable estimate therefore reflecting the urbanization process in Rio de Janeiro.

Considering these points, we infer that for most out-of-sample satellite images tested, the model exhibited a tendency to overpredict "vegetation". This outcome presents a limitation in maintaining temporal consistency across areas previously identified as "urban" or "other" which are unlikely to revert to native vegetation. We suggest that this behavior may be associated with noise in the Sentinel-2 Level-1C (L1C) imagery used for training, whereas the Level-2A (L2A) collection could have provided more temporally consistent visual patterns due to its atmospheric correction.

5. Projections

The number of samples within the time series is insufficient. Therefore it is not possible to run more sophisticated statistical models such as autoregressive models like ARIMA or state-space models, or deep learning models like LSTM. The best possible estimation for 2025, 2026 and 2027 would be a moving average model where:

$$prediction_{2025} = \frac{result_{2022} + result_{2023} + result_{2024}}{3} \tag{2}$$

$$prediction_{2026} = \frac{result_{2023} + result_{2024} + prediction_{2025}}{3} \tag{3}$$

This procedure would be repeated for 2027. This calculation gives the following results:

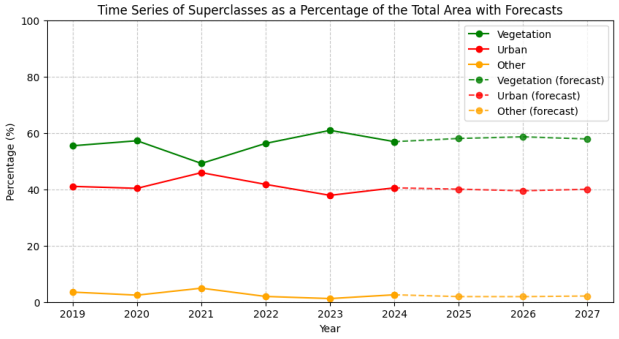

Figure 6. Percentage of the city covered in each superclass, with forecast.

6. Conclusion

Our model has good metrics and could capture the vegetation and urban coverage classes adequately in the test sample. The model reaches accuracy of almost 90%, however, some changes should be implemented to achieve better results and better match the baseline. At times the model overestimates the "vegetation" class at the expense of both of the other classes which creates further challenges in maintaining temporal consistency. This

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-W4-2025-197-2026

over representation of the ”vegetation” areas in the results indicates that our model still needs further improvements in order to fully map the UHI effect. With these considerations in mind, the model can be used, however, further development of the model is required. This includes the use of different methods to balance out the under-representation of the class ”others” and to further investigate the out-of-sample overestimation of the ”vegetation” class. Additionally, further developments could use the L2A collection. Even though this would result in the loss of the 2016 images, data augmentation techniques could be used to increase the number of samples and still guarantee good generalization of the model. Furthermore, data augmentation could also be used to increase the representation of the ”others” class. Another possibility could be to model the inner-classes such as ”mining”, and ”farming areas” individually to capture the full detail coverage Data Rio can provide.

References

Balaban, O., 2012. The negative effects of construction boom on urban planning and environment in Turkey: Unraveling the role of the public sector. *Habitat International*.

Campos-Taberner, M., García-Haro, F. J., Martínez, B., Moreno, L., Pérez-Suay, A., Camps-Valls, G., Gilabert, M. A., 2020. Understanding deep learning in land use classification based on Sentinel-2 time series. *Scientific Reports*, 10, 17188. https://doi.org/10.1038/s41598-020-74215-5.

Copernicus Data Space Ecosystem, 2025. Sentinel-2 mission data. `https://documentation.dataspace.copernicus.eu/Data/SentinelMissions/Sentinel2.html`. Accessed: 2025-09-20.

Earth Engine Data Catalog, 2025. Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-1C (TOA). Google Earth Engine.

Environmental Protection Agency, 2008. Reducing urban heat islands: Compendium of strategies - urban heat island basics.

European Space Agency, 2025. Sentinel-2 toolbox features. `https://step.esa.int/main/toolboxes/sentinel-2-toolbox/sentinel-2-toolbox-features/`. Accessed: 2025-09-20.

European Space Agency / Copernicus SentiWiki, 2025. S2 mission – spectral resolution. `https://sentiwiki.copernicus.eu/web/s2-mission#S2Mission-SpectralResolutionS2-Mission-Spectral-Resolution`. Accessed: 2025-09-20.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Jadon, S., 2020. A survey of loss functions for semantic segmentation. *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, IEEE.

Louzada, R. O., 2025. Taquari River Basin: a basin of contrasts, environmental challenges, and ecosystem opportunities. PhD thesis, Universidade Federal de Mato Grosso do Sul.

MapBiomas, 2025. Mapbiomas 10m. https://plataforma.brasil.mapbiomas.org/cobertura_10m.

Prefeitura da Cidade do Rio de Janeiro, 2016a. Cobertura vegetal e uso da terra 2016. `https://www.data.rio/datasets/a22768dc645043f5bdfd1affab888ded_0/explore`.

Prefeitura da Cidade do Rio de Janeiro, 2016b. O rio visto do espaço: Ilhas de calor urbano e mudanças climáticas. `https://www.data.rio/apps/5d9b36b0c4054369b39eb7bf6c2159d7/explore`.

Prefeitura da Cidade do Rio de Janeiro, 2018. Cobertura vegetal e uso da terra 2018. `https://www.data.rio/datasets/c32974e0db954842b7af9a4816d7a821_0/explore`.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, Springer, 234–241.

UN, 2025. 68% of the world population projected to live in urban areas by 2050, says un. https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html.

World Bank Group, 2023. Unlivable: What the urban heat island effect means for east asia's cities.

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., Atkinson, P. M., 2019. Joint Deep Learning for land cover and land use classification. *Remote Sensing of Environment*, 221, 173-187. `https://www.sciencedirect.com/science/article/pii/S0034425718305236`.

Zhenqi, H. (ed.), 2017. *Land Reclamation in Ecological Fragile Areas*. CRC Press.