Niterói, Universidade Federal Fluminense
ISSN 15177793 / 26748126 (eletrônico)
GEOgraphia, vol: 28, n. 60, 2026
DOI: 10.22409/GEOgraphia2026.v28i60.a50525

A GEOBIA AND LOGISTIC REGRESSION APPROACH FOR IDENTIFYING PRECARIOUS SETTLEMENTS USING VERY-HIGH-RESOLUTION IMAGERY: A CASE STUDY IN THE GRANDE ABC REGION, SÃO PAULO, BRAZIL

Bruno Dias dos Santos[1]
McMaster University
Hamilton, Ontario, Canada

Bruno de Medeiros Garibaldi Pereira[2]
Universidade de São Paulo (USP)
São Paulo, SP, Brasil

Carolina Moutinho Duque de Pinho[3]
Universidade Federal do ABC (UFABC)
Santo André, SP, Brasil

Enviado em 20 jun 2021 | Aceito em 22 jun. 2025

Abstract: This study proposes a methodology to identify precarious settlements using a Geographic Object-Based Image Analysis (GEOBIA) approach combined with a logistic regression model. The study area comprises 25 km² located between São Bernardo do Campo and Santo André in the Grande ABC Region, São Paulo State, Brazil, containing over 80 mapped precarious settlements with diverse morphological characteristics. Using exclusively open-source software, a Worldview-3 image, and a Digital Surface Model, we generated geographic objects at two hierarchical segmentation levels. The first level of objects was classified using a decision tree generated by the C4.5 machine learning algorithm into nine land cover classes. The second level consisted of a regular cellular grid, classified using a logistic regression model based on variables derived from sub-object information, neighborhood context, and cell-level metrics. Results showed Kappa values of 0.68 for the development area and 0.72 for the validation area, demonstrating that the methodology has potential for identifying precarious settlements from very-high-resolution satellite imagery.
Keywords: slums, precarious, informal, classification, model, remote sensing.

1. PhD student in Transportation Geography at McMaster University, Canada, specializing in spatial models to identify urban inequalities and inform public policy. I hold degrees in Environmental and Urban Engineering (Federal University of ABC, Brazil) and a Master's in Remote Sensing (INPE, Brazil). E-mail: brdiasdossantos@gmail.com. Orcid: https://orcid.org/0000-0002-6748-2038.
2. Bachelor of Statistics (2005) at University of São Paulo. E-mail: brunom@ime.usp.br. Orcid: https://orcid.org/0000-0002-2414-0855.
3. Professor at the Federal University of ABC (UFABC). Bachelor's degree in Geography from the Federal University of Rio de Janeiro (2000), a master's degree (2005) and a doctorate degree (2012) in Remote Sensing from the National Institute for Space Research (2005), and worked as a postdoctoral researcher in Public Administration from Getúlio Vargas Foundation (2013). E-mail: carolina.pinho@ufabc.edu.br. Orcid: https://orcid.org/0000-0002-7054-4463.

GEOBIA E REGRESSÃO LOGÍSTICA PARA IDENTIFICAÇÃO DE ASSENTAMENTOS PRECÁRIOS UTILIZANDO IMAGENS DE ALTÍSSIMA RESOLUÇÃO ESPACIAL: UM ESTUDO DE CASO NA REGIÃO DO GRANDE ABC, SÃO PAULO, BRASIL

Resumo: Este estudo propõe uma metodologia para identificar assentamentos precários utilizando Análise de Imagens Baseada em Objetos Geográficos (GEOBIA) combinada com um modelo de regressão logística. A área de estudo compreende 25 km² localizados entre São Bernardo do Campo e Santo André na Região do Grande ABC, Estado de São Paulo, Brasil, contendo mais de 80 assentamentos precários mapeados com características morfológicas diversas. Utilizando exclusivamente softwares de código aberto, uma imagem Worldview-3 e um Modelo Digital de Superfície, geramos objetos geográficos em dois níveis hierárquicos de segmentação. O primeiro nível de objetos foi classificado usando uma árvore de decisão gerada pelo algoritmo de aprendizado de máquina C4.5 em nove classes de cobertura do solo. O segundo nível consistiu em uma grade celular regular, classificada usando um modelo de regressão logística baseado em variáveis derivadas de informações dos sub-objetos, contexto de vizinhança e métricas ao nível das células. Os resultados mostraram valores de Kappa de 0,68 para a área de desenvolvimento e 0,72 para a área de validação, demonstrando que a metodologia tem potencial para identificar assentamentos precários a partir de imagens de satélite de altíssima resolução espacial.
Palavras-chave: favelas, precário, informaL, classificação , modelo, sensoriamento remoto

UN ENFOQUE GEOBIA Y REGRESIÓN LOGÍSTICA PARA LA IDENTIFICACIÓN DE ASENTAMIENTOS PRECARIOS UTILIZANDO IMÁGENES DE MUY ALTA RESOLUCIÓN ESPACIAL: UN ESTUDIO DE CASO EN LA REGIÓN DEL GRAN ABC, SÃO PAULO, BRASIL

Resumen: Este estudio propone una metodología para identificar asentamientos precarios utilizando Análisis de Imágenes Basado en Objetos Geográficos (GEOBIA) combinado con un modelo de regresión logística. El área de estudio comprende 25 km² ubicados entre São Bernardo do Campo y Santo André en la Región del Gran ABC, Estado de São Paulo, Brasil, conteniendo más de 80 asentamientos precarios mapeados con características morfológicas diversas. Utilizando exclusivamente software de código abierto, una imagen Worldview-3 y un Modelo Digital de Superficie, generamos objetos geográficos en dos niveles jerárquicos de segmentación. El primer nivel de objetos fue clasificado usando un árbol de decisión generado por el algoritmo de aprendizaje automático C4.5 en nueve clases de cobertura del suelo. El segundo nivel consistió en una cuadrícula celular regular, clasificada usando un modelo de regresión logística basado en variables derivadas de información de los sub-objetos, contexto de vecindad y métricas a nivel de las celdas. Los resultados mostraron valores de Kappa de 0,68 para el área de desarrollo y 0,72 para el área de validación, demostrando que la metodología tiene potencial para identificar asentamientos precarios a partir de imágenes satelitales de muy alta resolución espacial.
Palabras clave: barrios marginales, asentamientos precarios, asentamientos informales, clasificación, modelo, teledetección.

*

Introduction

One in eight people around the world lives in a slum (UN-Habitat, 2016). In developing countries, the number of urban residents living in slums increased by 28% in absolute value, from 689 million in 1990 to 1.1 million people in 2014, representing 30 percent of the urban population in developing countries.

In Brazil, for public policy purposes, slums are included within a larger category of settlements, named precarious settlements: segments of urban territory with varying magnitudes and types, but predominantly formed by residential areas occupied by low-income residents. These residents use multiple strategies to produce solutions for their housing needs autonomously, since neither the State nor the formal market meet their demands. In general, these settlements encompass numerous types, with emphasis on tenements (*cortiços*), slums, irregular low-income settlements, and housing estates produced by the government that are degraded.

There are some methodological approaches that use geoprocessing tools to identify precarious areas in Brazil, mainly to support the production of public housing policies. In general, these initiatives are based on census information, like the projects promoted by IBGE (2010) and CEM (2007; 2013). However, there are some limitations in the strictly use of census data, such as the intensive and time-consuming collection of data that requires a large financial investment, in addition to the time gap between the two censuses (in Brazil, one census is conducted every 10 years). Mahabir et al. (2020) raised three other limitations regarding census data for the identification of slums: census statistics are generally provided at the aggregate level of the city or neighborhood, failing to capture all the heterogeneity present in slums (DUQUE et al., 2012; GONÇALVES, 2018); often, slum residents are afraid of giving personal information and facing the misuse of these data, facilitating evictions by the authorities (GALEON, 2008; KAREKEZI, 2008; ALVES et al., 2011); and finally, even though there are updated data, the poor quality of the information can disable the use of these data to support and develop the public policies necessary to reduce the populations of slums (HENDERSON et al., 2012).

More recently, CDHU/UFABC (2018) and Feitosa et al. (2021) through the Methodology for Identification and Characterization of Precarious Settlements in Metropolitan Regions of São Paulo (MAPPA), identified the settlements of the Metropolitan Region of Baixada Santista, standing out by using a base of cellular representation and non-census variables. The MAPPA Project, in addition to identifying precarious areas, also classified them into urban occupation typologies according to their physical characteristics.

For the Grande ABC Paulista, Dos Santos, Pinho and Barboza (2019) investigated the potential of imagery from the Rapideye Satellite (five meters of spatial resolution) and Digital Elevation Models (DEM) to identify precarious settlements. The authors employed a Geographic Object-Based Image Analysis (GEOBIA) approach, with a segmentation based on a cellular grid of 100 x 100 meters and, despite not having obtained great results in terms of classification accuracy, it was one of the first studies to use data exclusively from remote sensing to identify precarious settlements of ABC and managed to work on a regional scale. More recently, Dos Santos et al. (2022) and Dos Santos et al. (2023) identified different urban settlements in cities across the Brazilian Amazon using open-source CBERS-4A imagery.

Overall, methodologies applied in Brazilian study sites for identifying precarious settlements have limitations such as restrictions to the limit of census sectors, general exclusions of physical characteristics of these areas, and dependency on data coming from visual interpretation (GONÇALVES, 2018; CDHU/UFABC, 2018). For this reason, studies using satellite images can be an alternative for identifying precarious areas (MAHABIR et al., 2020; KUFFER et al., 2016).

Hofmann (2001) proposed one of the first studies using remote sensing data for slums identification. For a region of Cape Town, South Africa, with an Ikonos image, he generated objects after multi-resolution and hierarchically linked segmentation at different spatial scales. The settlements were the super objects, which had several attributes of form and context, described after applying fuzzy functions. Subsequently, Hoffman, Strobl and Kux (2008) carried out studies for mapping slums in Rio de Janeiro, and after them, some other studies have investigated the potential of using remote sensing data to identify precarious settlements in Brazil (NOVACK and KUX, 2010; REUß, 2017; WURM and TAUBENBÖCK 2017; FRIESEN et al., 2019; MÜLLER et al. 2020).

The general objective of this paper is to identify precarious settlements using data obtained strictly from remote sensing. We applied GEOBIA techniques on Very-High-Resolution (VHR) imagery (0.3 meters of spatial resolution) combined with a logistic regression model to classify geographic objects, in a 25 km² study area located between São Bernardo do Campo and Santo André, São Paulo State.

The main contribution of this study is to automate the process to identify areas with a high chance of being precarious settlements, reducing the work for visual interpretation by technicians and guiding the work of field teams. To increase reproducibility, only open-source software was employed.

Materials and Methods

The methodology is summarized in five stages: selection of the study area, data preparation, classification of the land cover (first-level classification), identification of precarious settlements (second-level classification), and results evaluation. The following datasets and software were used (Table 1):

Table 1 - Datasets and softwares

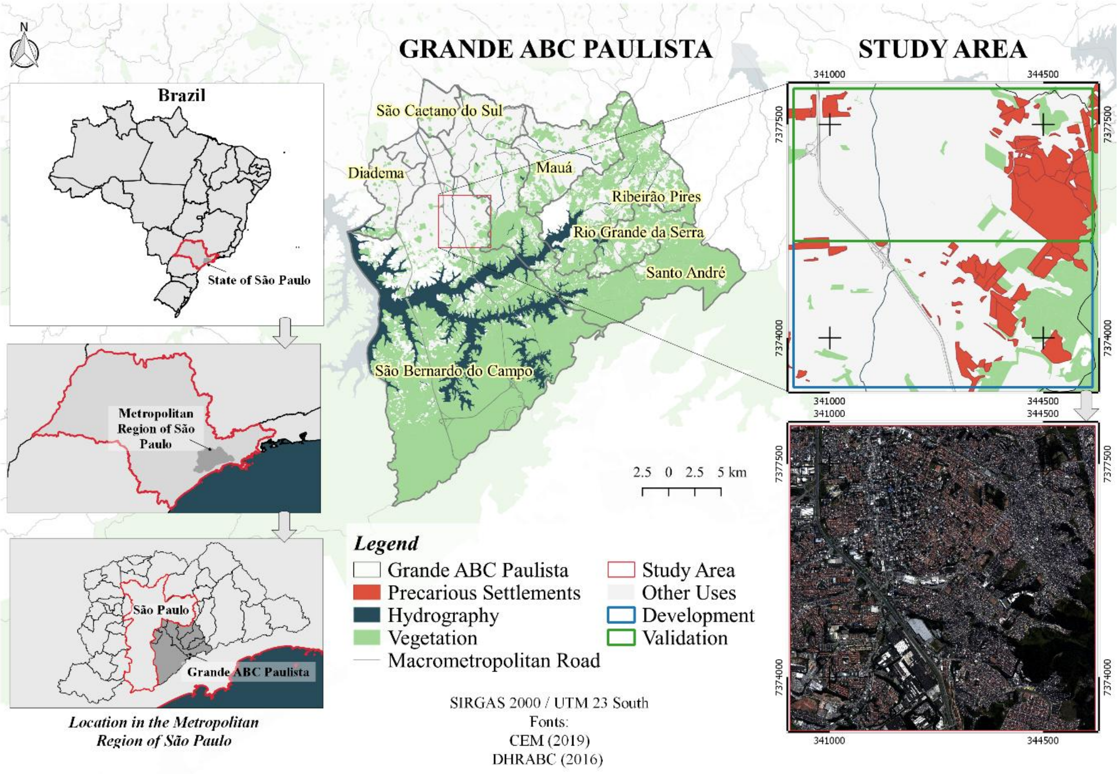
Product	Description
Raster files	Two Worldview 3 images, both acquired on 08/30/2015, one panchromatic with a spectral resolution of 0.3 m and the other multispectral with 0.6 m resolution and 8 bands: band 1 (coastal blue), band 2 (blue), band 3 (green), band 4 (yellow), band 5 (red), band 6 (red edge), band 7 (near infrared 1) and band 8 (near infrared 2), dated August 30, 2015, provided by Digitalglobe for research purposes only.
	One Digital Elevation Model from 2010/2011, with 5 meters of spatial resolution (scale of 1: 25,000) from Emplasa.
Vector files	Other vector files were used for the generation of thematic maps, such as the shapefile of the limits of the municipalities and the precarious settlements, among others.
Software	Qqis 3.10 for image pre-processing, segmentation, attribute extraction and thematic maps elaboration.
	Terraview 5.3.3 - Geodma Plugin for extracting attributes and elaborating the decision tree for classifying objects of the first level.
	RStudio (R) for developing the Logistic Regression Model.

Study Area

The study area is located in the Grande ABC Region, in the southeast of the São Paulo Metropolitan Region (RMSP). The Grande ABC is home to more than 2.7 million people in a territorial area of 828 km², concentrating almost 800 precarious settlements with almost 200,000 households, which represents about 25% of the total households.

More specifically, our study site refers to a 25km² area located between São Bernardo do Campo and Santo André (Figure 1), centered in the region known as *Montanhão* (Big Mountain) (DHRABC, 2016). There are 84 precarious settlements mapped in the study area.

Figure 1 - Study área



The study area was divided into two: (a) the development area, where the information from that region was used in the methodology; and (b) the validation area, where the methodology was applied to evaluate the applicability in other areas, maintaining the same conditions.

Data Preparation

First, we merged the multispectral and panchromatic images by applying the cubic pansharpening algorithm. Image fusion is a process that combines images with different spatial and spectral resolutions in order to obtain a final synthetic product with higher spatial and spectral resolutions. Next, we calculated the slope (in percentage) from the Digital Surface Model (5 meters of spatial resolution). Subsequently, the merged image went through band algebra techniques to create the indexes described in Table 2.

Table 2 - Indices created

Index	Description
NDVI $\frac{(NIR - Red)}{(NIR + Red)}$	The Normalized Difference Vegetation Index (NDVI) consists of the division of the near infrared (NIR) band by the red band, in a normalized manner. This index is used to analyze the presence and condition of natural or agricultural vegetation in images generated by remote sensors and is often used to measure the intensity of chlorophyll activity.
NDRI $\frac{(Red - Blue)}{(Red + Blue)}$	The Normalized Difference Roof Index (NDRI) consists of dividing the band corresponding to red by blue in a normalized way. Through this index, it is possible to identify the presence of ceramic tiles on buildings and areas with exposed soil, so that these areas have values closer to 1.
BAi $\frac{(Blue - NIR)}{(Blue + NIR)}$	The Bare Soil Area Index (BAI) is the normalized division of the blue band by that of the infrared, and is used to identify exposed soil, helping to separate these areas from areas of undergrowth and fields.

We adapted the Generic Slums Ontology (GSO) created by Kohli et al. (2021) and the interpretation key of Santos and Pinho (2017) to create an ontology for the precarious settlements of our study area (Table 3). The GSO highlights relevant indicators regarding the shape of the settlements, size of residences, roofing material, road regularity, vegetation, and the location of the settlements at three levels of analysis: environment, settlement, and object. This was posteriorly translated into elements of photointerpretation such as color, size, shape, location, and texture and guided the stages of segmentation and classification.

Table 3 - Ontology for the precarious settlements of the study área

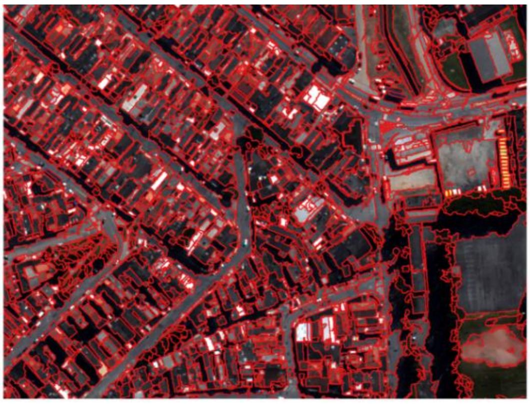
Level	Indicators	Local Obsevation	Parameterisation
Environment	Location	Usually located in areas of medium and high slopes, closer to major highways and power lines	Slope and elevation by the DEM and extraction of roads by classification of the satellite image
	Neighbourhood characteristics	Close to middle socioeconomic-status neighborhoods and other slums	Socioeconomic information was not used, but comparison with the surrounding objects was made by creating an index of difference of neighbors
Settlement	Shape	Large and irregular settlements	Not used
	Density	Highly dense buildings, low vegetation and open spaces	Texture, contrast, NDVI and area of roof and vegetation
Object	Building	Variable format, 10 to 35 m² in area and predominantly fiber cement roofs;	Area and size of segments, spectral and shape attributes
	Access network	Irregular shape, paved access, generally narrow streets	Extraction of roads by classification of the satellite image

Classifying the land cover

After defining the GSO for our study area, we segmented the processed image to create objects with homogenous pixel characteristics. We adopted a bottom-up segmentation strategy, first creating subobjects for a finer and more detailed level (first level) that were later grouped into a cellular grid as superobjects, which are hierarchically higher (second level). This hierarchical organization stored association relationships between superobjects and their respective subobjects, as well as bringing neighborhood information to objects of the same level (PINHO, 2005).

The first level of objects was created after segmenting the pansharpened imagery, using the Mean-Shift Segmentation clustering algorithm of the Orfeo ToolBox Plugin (OTB) (OTB, 2021). This algorithm first defines a window around each point and averages the gray tones of the image; then, it changes the center of the windows for each centroid and repeats the procedure until iteratively converges. The result segments the image into several objects (segments). Our choice for this algorithm was motivated because it presented a better performance in processing time and segmentation of the objects compared to the other four algorithms available in the OTB Plugin. The size and shape of segments are controlled by selecting values of spatial radius, range radius, maximum number of iteractions, and minimum region size[4]. In this study, we adopted the following values to perform our segmentation (Figure 2): a) spatial radius: 5; b) range radius: 110; c) maximum number of iteractions: 100; and d) minimum region size: with a default value of 100.

Figure 2 - Segmentation result



The idea was to generate geographic objects in sizes and shapes that would later be classified among the classes: fiber cement roofing, ceramic roofing, high gloss roofing, exposed soil, asphalt, dense vegetation, sparse vegetation, logistic warehouses, and shadow. These classes were defined after the visual analysis of the pansharpened image and analysis of the interpretation key to identify the main targets that make up the objects within the study area.

For each segment, we extracted the mean, minimum, maximum, and standard deviation values of NDVI, NDRI, BAi, DEM, and declivity. Additionally, we also extracted spectral attributes (such as average, minimum, maximum, variance, and texture) and shape attributes (such as area, perimeter, and degree of rectangle) for the objects using the GeoDMA software (INPE, 2011). Spectral

[4] For parameter definitions, see (OTB, 2021).

information was extracted only from the bands referring to the blue, green, red, near infrared, and infrared (NIR), as these are common bands and can be found in images from other satellites[5]. After extracting the attributes, about 900 samples of segments were collected.

We used the C4.5 decision tree algorithm (QUILAN, 1993) to classify the segments. In this algorithm, the choice of attributes is made by an entropy metric called information gain, so that the attributes with greater information gains are selected in the construction of the decision tree. The information is organized and distributed in nodes, descending branches, and leaves that are associated with each class. The first attribute at the top of the tree represents the root, and the lines represent the branches connected to the nodes related to some other attribute, moving to the right or to the left depending on a threshold value. Finally, the leaves at the ends of the branches represent the classes (JUSTINO, 2014).

Classifying precarious settlements

The second classification level is formed by a grid of regular cells with dimensions of 100 by 100 meters (2,600 cells in total). This cell size was adopted because this is the same size as the length adopted in the MAPPA methodology (CDHU/UFABC, 2018). In addition, the average slum area for the study region is 51,000 square meters.

Next, we created a hierarchical relationship between the grid (hierarchically higher, superobjects), and the segments of the first level (hierarchically lower, subobjects). Then, for each cell we calculated the percentage area coverage of its sub-objects and extracted the same attributes previously obtained for the segments, but for the cell grid. Following, we created attributes based on neighborhood relationships as shown in Table 4. At the end of this process, over 400 attributes were obtained for each cell.

Table 3 - Neighbourhood variables created for the second level

Index	Description
$\underline{\boldsymbol{Y}} = \frac{\sum_{i=1}^{8} \boldsymbol{X_i}}{\boldsymbol{8}}$	Since $X_1, X_2, X_3, X_4, X_5, X_6, X_7$ *and* X_8 are the neighbouring and adjacent cells of Cell Y, the Mean Neighbourhood Value $\underline{Y}$ was calculated for all variables in the 8 adjacent cells using Equation 5.
$\boldsymbol{Ind_{viz}} = \frac{(\boldsymbol{Y} - \underline{\boldsymbol{Y}})}{(\boldsymbol{Y} + \underline{\boldsymbol{Y}})}$	The Normalized Difference Index of the Neighborhood varies from -1 to 1, where -1 indicates the cell has an attribute (Y) much smaller than the average of that same attribute for the bordering cells, and 1 indicates that the cell has an attribute (Y) much greater than the mean of this same attribute for the bordering cells.

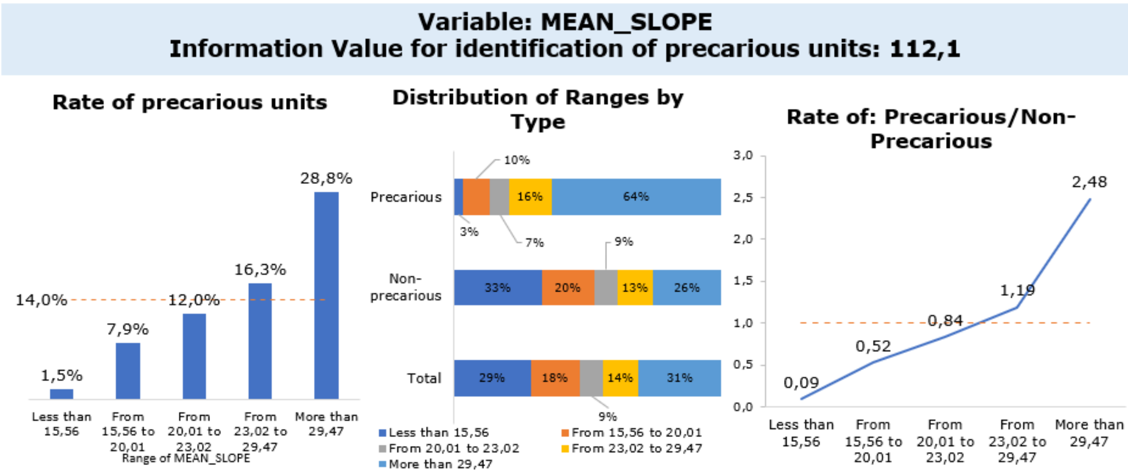
The continuous variables were then categorized, on a process that divided their continuum scale into a small number of intervals with discrete limits (KERBER, 1992). These categories were defined by maximizing the information value – a measure used to reduce independent variables that have little relation to the dependent variable and can be used to compare the predictive power among all independent variables.

Figure 3 shows that 64% of the cells that are precarious have an average slope greater (*MEAN_SLOPE*) than 29.5%, while only 26% of the non-precarious cells have an average slope in this category. A precarious cell is 2.48 times more likely to have an average slope above 29.5% than a non-

[5] For the complete list of attributes that can be extracted by GeoDMA, see (INPE, 2011).

precarious cell, which makes the average slope a potentially good variable for the separation between precarious and non-precarious areas.

Figure 3 - Categorization of mean slope Variable



Only variables with an information value higher than 0.5 and correlations lower than 0.65, analyzed two by two, were kept for modeling. The eliminations of variables aimed to reduce the redundancy of information, since highly correlated variables tend to measure the same information and variables with low information value have low predictive power. At the end of the process, only 54 variables remained.

Logistic model

We performed a logistic regression to model the relationship between the variables obtained from remote sensing and precarious settlements. A logistic regression model expresses the probability of an event occurring based on a set of n independent variables (STRAUSS, 1992) (Equation 6):

$$P(Y = 1 \mid X_1 = x_1, \dots, X_n = x_n) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)}} \quad (6)$$

The coefficients $\beta_0, \beta_1, \dots \beta_n$ are estimated from a set of data. Probabilities close to 1 indicate that the state of Y = 1 (precarious) is very probable, and probabilities close to 0 indicate otherwise (not precarious). For the construction of the statistical model, only those cells that had more than 25% of their area overlaid on a precarious settlement defined by DHRABC (2016) were determined as precarious cells. This decision was made to exclude cells that could "confuse" the model and decrease product quality.

During the modelling, the backward stepwise elimination method was used to select variables, starting with all independent variables. At each step, the model drops variables one by one, using a model performance comparison criterion to identify variables that least improve the model and eliminates them at the end of each step, then repeats this procedure until there is no further improvement in the model (SILVA, 2016).

Results

The result of the classification of the first level of objects is shown in Figure 4, and the confusion matrix, calculated by object on a simple random sample, is shown in Table 5. The confusion matrix shows that all classes present high percentages of user accuracy, with a highlight for fiber cement roofing and high gloss roofing, both with 100% accuracy rates. Even the class with the lowest user accuracy and kappa per class, logistics warehouses, presented indices higher than 90%.

Figure 4 - Land cover classification (first level)

In terms of producer accuracy, the classes warehouses, shadow, dense vegetation, and high gloss coverage presented 100% hit rates, and the sparse vegetation and exposed soil classes had the lowest rates (96%). Both the global accuracy and the kappa of the land cover classification are 0.98. Despite this, the parameters of the decision tree were not applied in a validation area due to the unavailability of a second image, which means that the quality of the classification was only attested for the region of the study area.

Table 5 - Confusion matrix of the first level classification

Classification	Reference Data: Fiber cement	High Gloss	Ceramic	Warehouses	Sparse Veg.	Exp. Soil	Shadow	Dense Veg.	Asphalt	N	User Accuracy	Class Kappa
Fiber cement	**142**									142	1.00	1.00
High Gloss		**209**								209	1.00	1.00
Ceramic			**133**		1	2				136	0.98	0.97
Warehouses	4			**58**					1	63	0.92	0.92
Sparse Veg.			1		**87**	1				89	0.98	0.98
Exp. Soil					3	**75**				78	0.96	0.96
Shadow			1				**29**			30	0.97	0.97
Dense Veg.								**30**	1	31	0.97	0.97
Asphalt		1							**122**	123	0.99	0.99
N	146	210	135	58	91	78	29	30	124	901		
Producer Accuracy	0.97	1.00	0.99	1.00	0.96	0.96	1.00	1.00	0.98			

Kappa of Classification	0.98
Global Accuracy	0.98

The estimated logistic model is shown in Table 6. Generally, the variables that remained in the model after the stepwise process translated the elements of image interpretation described in our ontology. The model shows that areas with higher elevation and slopes have higher probability of being precarious, aligning with other studies that connected the precarious settlements of the Grande ABC with environmental risk (DHRABC, 2016; DENALDI and CARDOSO, 2018). The model also shows that precarious areas are often surrounded by high built-up density and have low vegetation, without the presence of a structuring road network, and with roofs mainly built with fiber cement. Variables related to spectral information also seem to have a higher predictive power to distinguish precarious from non-precarious areas, especially those linked to the NIR and Blue bands. These two bands assist in the separation of fiber cement and ceramic roofs.

Table 5 - Logistic regression model to identify precarious settlements in the study área

Variable	*Description*	*Category*	*B*	*Error*	*Pr(>|z|)*
(Intercept)	-	-	-2.7906	0.9802	0.004415
VIZ_MEAN_SLOPE	Average slope value of neighboring cells	> 24.37% e <= 29.60%	1.087	0.3472	0.001745
		> 29.60%	3.189	0.3608	< 2.00E-16
VIZ_ASBESTOS	Average value of (%) of area classified as "fiber cement coverings" in neighboring cells	> 18.38% e <= 20.57%	1.6654	0.6546	0.010947
		> 20.57%	2.2853	0.4464	0.000000306
MAX_SLOPE	Maximum cell slope value	> 42.38%	1.1818	0.4719	0.012258
VEGETATION	(%) of area classified as "Dense Vegetation" in the cell	> 54.60 % e <= 73.25%	-1.6933	0.6353	0.007694
		> 73.25	-3.3888	1.1326	0.002771
BRATIO_B0	It describes the contribution of blue band to the region. Calculated by the value of the Blue band divided by the sum of the values of all other bands	> 0.21	1.3256	0.6387	0.037939

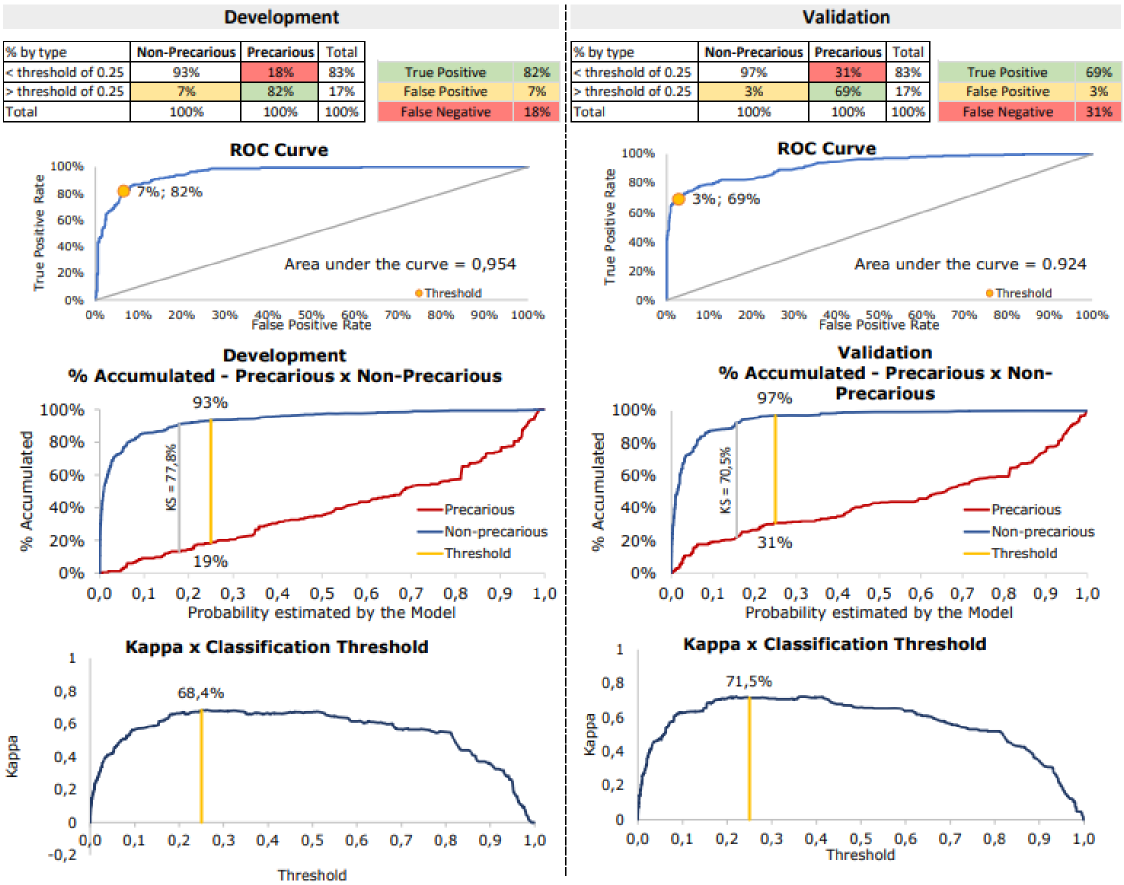
ASPHALT	(%) of area classified as "Asphalt" in the cell	> 17.02 e <= 21.60%	-1.0282	0.3821	0.007135
		> 21.60%	-1.3016	0.3882	0.0008
VIZ_CONTSE_B5	Average contrast value of neighbouring cells for the Near Red band	> 11.49	-3.0675	1.1311	0.006691
VIZ_MAX_BAI	Mean value of the BAI Index maxima for neighboring cells	> 65.07%	-1.0119	0.3044	0.000888
BRATIO_B4	It describes the contribution of the NIR band to the region. Calculated by the value of the NIR band divided by the sum of the values of all other bands	> 0.22 e <= 0.32	-1.3106	0.358	0.000251
		> 0.32	-3.1774	0.7661	0.0000336
VIZ_VEGETATION	Average value of (%) of area classified as "Dense Vegetation" in neighboring cells	> 4.15%	-1.1229	0.3224	0.000495
IND_VIZ_MIN_MDS	Normalized difference between the mean of the neighbors' minimum DEM values and the cell value	> -0.0048 e <= -0.0035	-1.043	0.4863	0.031966
		<= -0.0035	-1.4638	0.3832	0.000133
ENTRSE_B6	Measures the uniformity of the NIR Band	> 4.74	2.4702	0.5907	0.0000289
IND_VIZ_MODE_B5	Normalized Neighborhood Difference Index for the Near Red Band's Mode attribute	> 0.034	-0.7267	0.2951	0.013806
IND_VIZ_MAX_MDS	Normalized difference between the average of the maximum MDS values of the neighbors and the cell value	> 0.005	1.1229	0.4016	0.005172
VIZ_MODE_B5	Average of Mode for the Near Red Band of neighboring cells	> 91.36	-0.9558	0.4576	0.036754

We evaluated the regression model based on the Kolmogorov-Smirnov (KS) statistic, the Receiver Operating Characteristic (ROC) curve, and the Kappa. The KS measures the maximum separation between the cumulative relative frequency of 0 (not precarious) and 1 (precarious). For the development and validation areas, the KS obtained was 77.8% and 70.5%, respectively. Although the model lost some performance in validation, the indices are still similar.

The ROC curve is a graphical representation that illustrates the performance of a binary classifier system as its threshold of discrimination varies. It compares the frequency of true positives and false positives for each point on the curve. From the ROC curve, it is possible to calculate the area under the curve (AUC), obtained by the total area of an accumulated distribution independent of a threshold. The AUC value is an assessment of the overall performance of a test, ranging from 0 (where all predictions in the model are wrong) to 1 (where all predictions in the model are correct). Our model resulted in 0.95 AUC for the development area, and in 0.92 AUC for the validation area, demonstrating that the proposed classification model is significantly better than a random model.

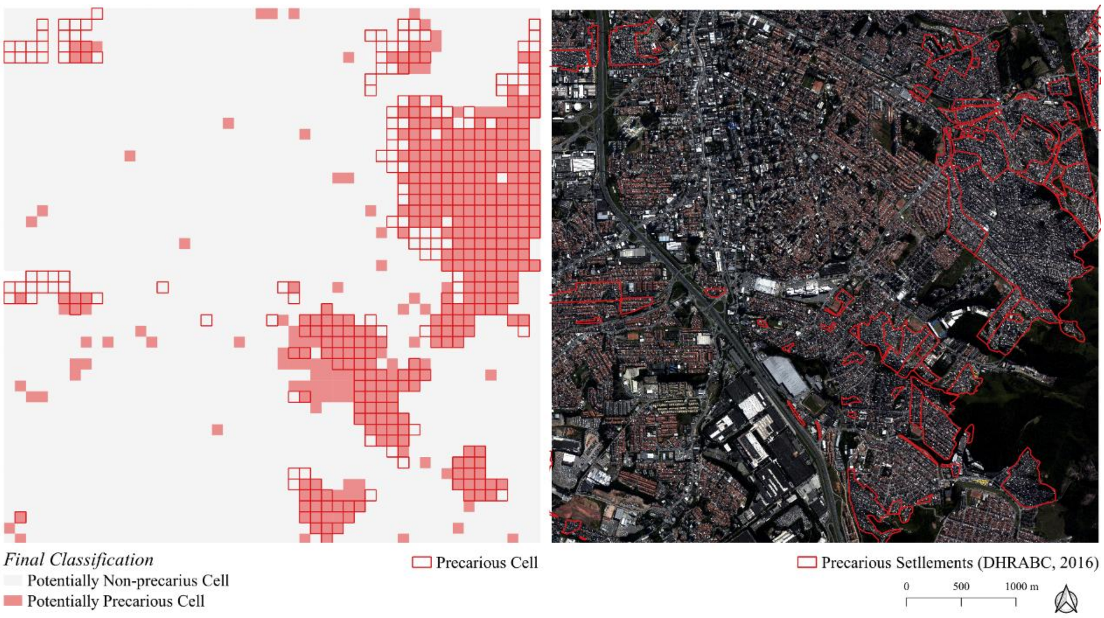
Like the ROC Curve, a Kappa index curve was generated to find the cutoff point that would maximize the true positive rate, bringing the highest quality to the classification for the development and validation areas. After the analysis of the ROC and Kappa curves (Figure 5), the visual interpretation of the Worldview 3 image, and the perimeters of precarious settlements, a threshold of 0.25 was adopted to separate the non-precarious and potentially precarious areas, generating the classification shown in Figure 6.

Figure 5 - (a) KS Statistics, ROC Curve and Kappa of the development and validation áreas



From a qualitative point of view, it is possible to verify that the classification was able to identify most of the precarious cells correctly. However, there were some commission errors (non-precarious areas were included as precarious) and omission errors (precarious areas not identified). As in the first level, a confusion matrix was calculated to evaluate the final classification, shown in Table 7. The 0.25 cutoff applied maximized the Kappa of the development area and returned Kappa of 0.68. For the validation area, the index was 0.72. Unlike the land cover (first classification level), at least for regions with urban characteristics similar to the study area, we have proof through the validation area that the model created can be applied in other regions without losing much of its performance.

Figure 6 - Classification of the second level



We believe that the construction of a second level composed of a network of regular cells and the generation of attributes for objects of this level, with the construction of attributes related to the neighbors, are the two main reasons for the high accuracy levels. However, for future work, the use of another metropolitan region for validation is indicated.

Table 7 - Confusion matrix of the second level classification

Development	Reference Data			User Accuracy	Class Kappa
Classification	Non-precarious	Precarious	N		
Non-precarious	**983**	30	1013	0.97	0.78
Precarious	69	**135**	204	0.66	0.61
N	1052	165	1217		
Precisão	0.93	0.82			

Kappa of Classification	0.6844
Global Accuracy	0.9187

Validation	Reference Data			User Accuracy	Class Kappa
Classification	Non-precarious	Precarious	N		
Non-precarious	**951**	80	1031	0.92	0.63
Precarious	28	**179**	207	0.86	0.83
N	979	259	1217		
Produtor	0.97	0.69			

Kappa of Classification	0.72
Global Accuracy	0.91

Although we consider that our methodology resulted in an overall good classification of precarious areas, we make some considerations for future work. Regarding the GEOBIA approach, it is known that it has some limitations. Mahabir et al. (2020) highlight some of them: (a) the parameters used for segmentation are interconnected and often subjectively defined; (b) very close residences are not well segmented; and (c) some targets may have similar spectral characteristics, reducing the segmentation performance. In addition, Kuffer et al. (2016) identified that GEOBIA approaches are those that showed the greatest variance in performance among the techniques most used for mapping slums, being highly linked to the complexity of the morphology. The author suggest the use of auxiliary data in the classification process and linking the products generated by satellite images to data of a socioeconomic nature. In addition to socioeconomic data, Mahabir et al. (2020) shed light on a range of Volunteered Geographic Information (VGI) data and Ambient Geospatial Information (AGI) that can be used to map and to better understand slums. Data concerning the characterization of the physical environment (such as proximity to rivers and wetlands), and data for morphological characterization of the settlements can also be applied to increase the classification accuracy.

We identified that a large number of objects of the first level were classified as shadow, which obstructed parts or entire residences in some of the cells. This problem was also discussed by Novack and Kux (2010) and by Mahabir et al. (2020) when classifying slums. When we look at the statistical model, we notice the lack of explanatory variables related to the type of roof coverage, which can be explained by the large number of segments classified as shadow in the areas of high declivity in the regions where the settlements are located.

Conclusion

The main objective of this study was to identify precarious settlements using only data extracted from remote sensing imagery, a technology whose use in Brazilian research has been quite incipient in the identification of slums. To achieve this objective, we combined GEOBIA with a logistic regression model for a study area located between São Bernardo do Campo and Santo André, in the Metropolitan Region of São Paulo, with a framework built exclusively with open-source free software.

We believe that the proposed methodology can contribute to the semi-automatic classification of satellite imagery for urban environments and for the identification of precarious areas. In our perspective, although the Worldview 3 image has enabled a good classification for the areas of development and validation (Kappa of 0.68 and 0.72, respectively), we believe that the proposed methodology has the potential to be applied to other free images, such as those from the CBERS- 04A, even though these may have a lower spatial resolution. That would mean great potential in the classification of images on a metropolitan scale. Another indication for future work is to validate the methodology using a similar study area. In this work, this was not possible due to the unavailability of a Worldview 3 image from another region, which corroborates the previous indication of using free images.

However, we emphasize that more than creating a “perfect model” for the identification of slums, there is the need to validate it with different stakeholders, such as public managers, residents, and civil society organizations, among other local actors, which makes the difference for real appropriation of the elaborated model. The mapping of precarious settlements carried out in loco or by visual interpretation is time consuming with high financial cost, requiring the mobilization of a large amount of specialized workforce. The presented methodology can be a complement to the works carried out by the municipalities to address such limitations. Currently, there are millions of people living in precarious settlements. With this study, we hope to contribute to the identification process of these precarious settlements by the government, leading to actions and interventions that improve the lives of the residents of these regions.

Referências

ALVES, M. C.; MORAIS, M.L.; ESCUDER, M.M.; GOLDBAUM, M.; BARROS, M.B.; CESAR, C.L.; CARANDINA, L. Household sampling in slums in surveys. Revista de Saúde Pública, v. 45, n. 6, p. 1099-1109, 2011. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102011000600012&lng=pt&nrm=iso. Acesso em: 21 mai. 2021.

ALVES, Mauro Correia. Estratégias para o desenvolvimento de modelos de credit score com inferência de rejeitados. 2008. 93 f. Tese (Doutorado) - Curso de Ciências, Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2008. Disponível em: https://www.teses.usp.br/teses/disponiveis/45/45133/tde-14102008-204609/publico/DissertacaoMauro.pdf. Acesso em: 12 jun. 2020.

BRASIL. Ministério das Cidades. (2010). Guia para mapeamento e caracterização de assentamentos precários. Brasília. Disponível em: <http://www.capacidades.gov.br/biblioteca/detalhar/id/181/titulo/guia-para-o-mapeamento-e-caracterizacao-de-assentamentos precarios#:~:text=A%20Secretaria%20Nacional%20de%20Habita%C3%A7%C3%A3,dados%20sobre%20os%20assentamentos%20prec%C3%A1rios>. Acesso em: 21 mai. 2021.

CÂMARA, F.P. Psiquiatria E Estatística V: Validação De Procedimentos Diagnóstica Pela Curva ROC. Psychiatry on line Brasil, v. 14, n. 4, 2009. Disponível em: http://www.polbr.med.br/ano09/cpc0409.php. Acesso em: 19 jun. 2021.

CDHU/UFABC. FEITOSA, F.F. et al. (2019). Relatório final: Metodologia para Identificação e Caracterização de Assentamentos Precários em Regiões Metropolitanas Paulistas (MAPPA). São Bernardo do Campo.

CENTRO DE ESTUDOS DA METRÓPOLE (CEM/CEBRAP). (2007). Assentamentos precários no Brasil urbano. Brasília: Ministério das Cidades. Disponível em: https://direito.mppr.mp.br/arquivos/File/relatorio-assentamentos-precatorios-no-brasil-urbano.pdf. Acesso em: 21 mai. 2021.

CENTRO DE ESTUDOS DA METRÓPOLE (CEM/CEBRAP). (2013). Diagnóstico dos assentamentos precários nos municípios da Macrometrópole Paulista. São Paulo: Cem. 2 v. (1). Disponível em: https://centrodametropole.fflch.usp.br/sites/centrodametropole.fflch.usp.br/files/user_files/ckeditor/655-Relatorio_I_Assentamentos_Fundap_Final_logo.pdf. Acesso em: 21 mai. 2021.

DA FONSECA FEITOSA, F.; VASCONCELOS, V. V.; DE PINHO, C. M. D.; DA SILVA, G. F. G.; DA SILVA GONÇALVES, G.; DANNA, L. C. C.; LISBOA, F. S. IMMerSe: An integrated methodology for mapping and classifying precarious settlements. Applied Geography, v. 133, p. 102494, 2021. DOI: https://doi.org/10.1016/j.apgeog.2021.102494.

DENALDI, R. (2009). Assentamentos precários: identificação, caracterização e tipologias de intervenção. In: DENALDI, R. (org.). Ações integradas de urbanização de assentamentos precários. Brasília: Ministério das Cidades. p. 93-128. Disponível em: http://www.capacidades.gov.br/media/doc/biblioteca/SNH003.pdf. Acesso em: 21 mai. 2021.

DHRABC. (org.). (2016). Diagnóstico habitacional regional do Grande ABC. Santo André: Ufabc. Disponível em: http://lepur.com.br/wp-content/uploads/2017/12/03-Diagnostico-Habitacional-Regional-do-Grande-ABC-Relat%C3%B3rio-Final.pdf. Acesso em: 19 jun. 2021.

DOS SANTOS, B. D. Identificação de assentamentos precários no Grande ABC a partir de dados de sensoriamento remoto. 2019. 84 f. TCC (Graduação em Engenharia Ambiental e Urbana), CECS, Universidade Federal do ABC, Santo André, 2019.

DOS SANTOS, B. D.; PINHO, C. M. D.; JESUS, T. B. (2019). Níveis de consolidação de assentamentos precários a partir de dados de sensoriamento remoto. In: XIX Simpósio Brasileiro de Sensoriamento Remoto, 19., 2019, Santos. Anais [...]. São José dos Campos: INPE. p. 3224-3227. Disponível em: https://proceedings.science/sbsr-2019/papers/niveis-de-consolidacao-de-assentamentos-precarios-a-partir-de-dados-de-sensoriamento-remoto. Acesso em: 22 mai. 2021.

DOS SANTOS, B. D.; DE PINHO, C. M. D.; OLIVEIRA, G. E. T.; KORTING, T. S.; ESCADA, M. I. S.; AMARAL, S. Identifying Precarious Settlements and Urban Fabric Typologies Based on GEOBIA and Data Mining in Brazilian Amazon Cities. Remote Sensing, v. 14, n. 3, p. 704, 2022. Disponível em: https://doi.org/10.3390/rs14030704.

DOS SANTOS, B. D.; DE PINHO, C. M. D.; OLIVEIRA, G. E. T.; KORTING, T. S.; ESCADA, M. I. S.; AMARAL, S. Identifying Precarious Settlements and Urban Fabric Typologies Based on GEOBIA and Data Mining in Brazilian Amazon Cities. Remote Sensing, v. 14, n. 3, p. 704, 2022. DOI: https://doi.org/10.3390/rs14030704.

DUQUE, J. C.; VICENTE, R.; NORENA, M. (2012). Stepwise Procedure to Determine a Suitable Scale for the Spatial Delimitation of Urban Slums. In: VÁZQUEZ, E. F.; MOROLLÓN, F. R. Defining the Spatial Scale in Modern Regional Analysis: new challenges from data at local level. Nova York: Springer. p. 237-254. Disponível em: https://www.springer.com/gp/book/9783642319938. Acesso em: 22 mai. 2021.

FRIESEN, J.; TAUNBENBÖCK, H.; WURM, M.; PELZ, P. F. Size distributions of slums across the globe using different data and classification methods. European Journal of Remote Sensing, v. 52, n. 2, p. 99-111, 2019. Disponível em: https://www.tandfonline.com/doi/full/10.1080/22797254.2019.1579617. Acesso em: 22 mai. 2021.

GALEON, F.A. (2008). Estimation of Population in Informal Settlement Communities Using High Resolution Satellite Image. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37., 2008, Beijing. Proceedings... Beijing: ISPRS. p.1377-1382. Disponível em: https://www.isprs.org/proceedings/XXXVII/congress/4_pdf/242.pdf. Acesso em: 21 mai. 2021.

GeoDMA 5.3.3. (2011). São José dos Campos, SP: Brazil's National Institute for Space Research (INPE). Available at http://www.dpi.inpe.br/geodma. Accessed in 08 september 2020.

GONÇALVES, G. S. Identificação de Assentamentos Precários na Região do Grande ABC: Uma abordagem estatística. 2018. 83 f. TCC (Graduação em Engenharia Ambiental e Urbana), CECS, Universidade Federal do ABC, Santo André, 2018.

HENDERSON, J. V.; STOREYGARD, A.; WEIK, D. N. Measuring Economic Growth from Outer Space. American Economic Review, v.102, n.2, p.994-1028, 2012. Disponível em: https://www.aeaweb.org/articles?id=10.1257/aer.102.2.994. Acesso em: 21 mai. 2021.

HOFMANN, P. (2001). Detecting Informal Settlements from Ikonos Image Data Using Methods of Object Oriented Image Analysis: an example from Cape Town (South Africa). In: Remote Sensing of Urban Areas, 1., 2001, Regensburg. Proceedings [...]. Regensburg: Geographische Schriften. p. 107-118. Disponível em: https://intranet.ifs.ifsuldeminas.edu.br/joao.tavares/Material_Cadastro_e_Loteamento/CadastroTecnicoMultifinalitario%20-%209EAC/Material%20auxiliar/kapstadt%5B1%5D.pdf. Acesso em: 19 jun. 2021.

HOFMANN, P.; STROBL, J.; BLASCHKE, T.; KUX, H. (2008). Detecting informal settlements from QuickBird data in Rio de Janeiro using an object based approach. In: BLASCHKE, T.; LANG, S.; HAY, G. J. (ed.). Object-Based Analysis: spatial concepts for knowledge-driven remote sensing applications. Berlim: Springer. p. 531-553. Disponível em: https://link.springer.com/chapter/10.1007/978-3-540-77058-9_29. Acesso em: 22 mai. 2021.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). (2010). Censo demográfico (2010): aglomerados subnormais: primeiros resultados. Rio de Janeiro: IBGE. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/periodicos/92/cd_2010_aglomerados_subnormais.pdf. Acesso em: 21 mai. 2021.

JUSTINO, Rodrigo Cesário. Mapeamento do uso da terra e da cobertura vegetal da subbacia hidrográfica do Rio São Tomé, Sul de Minas Gerais, por mineração de dados utilizando imagens IRS-P6/LISS III. 2014. 69 f. Dissertação (Mestrado) - Curso de Ecologia e Tecnologia Ambiental, Universidade Federal de Alfenas, Alfenas, 2014. Disponível em: https://www.unifal-mg.edu.br/ppgca/system/files/imce/Defesa/Dissertação_Rodrigo%20C.%20Justino.pdf. Acesso em: junho de 2020.

KAREKEZI, S.; KIMANI, J.; ONGORU, O. Energy Access among the Urban Poor in Kenya. Energy for Sustainable Development, v.12, n.4., p.38–48, 2008. Disponível em: https://www.sciencedirect.com/science/article/pii/S0973082609600065. Acesso em: 21 mai. 2021.

KERBER, R. (1992). Chimerge: discretization of numeric attributes. In: National Conference on Artificial Intelligence, 10., 1992, Palo Alto. Proceedings [...]. Palo Alto: AAAI. p. 123-128. Disponível em: https://www.aaai.org/Papers/AAAI/1992/AAAI92-019.pdf. Acesso em: 19 jun. 2021.
KOHLI, D.; SLIUZAS, R.; KERLE, N.; STEIN, A. An ontology of slums for image-based classification. Computers, Environment and Urban Systems, v.36, n.2, p.154-163, 2012. Disponível em: https://www.sciencedirect.com/science/article/pii/S0198971511001128. Acesso em: 22 mai. 2021.
KUFFER, M.; PFEFFER, K.; SLIUZAS, R. Slums from Space-15 Years of Slum Mapping Using Remote Sensing. Remote Sensing, v.8, n.6, p. 1-29, 2016. Disponível em: https://www.mdpi.com/2072-4292/8/6/455. Acesso em: 21 mai. 2021.
LANDIS, J. Richard; KOCH, Gary G. The measurement of observer agreement for categorical data. Biometrics, p. 159-174, 1977. Disponível em: https://www.jstor.org/stable/2529310. Acesso em: 19 jun. 2021.
LEONITA, G.; KUFFER, M.; SLIUZAS, R.; PERSELLO, C. Machine Learning-Based Slum Mapping in Support of Slum Upgrading Programs: The Case of Bandung City, Indonesia. Remote Sensing, v. 10, p. 1522, 2018.
LIN, A. Z. (2013). Variable Reduction in SAS by Using Weight of Evidence and Information Value. In: SAS Global Forum 2013, 1., 2013, Baltimore. Forum. Baltimore: SAS. p. 1-12. Disponível em: https://support.sas.com/resources/papers/proceedings13/095-2013.pdf. Acesso em: 19 jun. 2021.
MAHABIR, R.; CROITORU, A.; CROOKS, A. T.; AGOURIS, P.; STEFANIDIS, A. A Critical Review of High and Very High-Resolution Remote Sensing Approaches for Detecting and Mapping Slums: trends, challenges and emerging opportunities. Urban Science, v.8, n.2, p. 1-38, 2018. Disponível em: https://www.mdpi.com/2413-8851/2/1/8. Acesso em: 21 mai. 2021.
MÜLLER, I.; TAUBENBÖCK, H.; KUFFER, M.; WURM, M. Misperceptions of Predominant Slum Locations? Spatial Analysis of Slum Locations in Terms of Topography Based on Earth Observation Data. Remote Sensing, v. 12, n. 15, p. 2474-2493, 2020. Disponível em: https://www.mdpi.com/2072-4292/12/15/2474. Acesso em: 22 mai. 2021.
NOVACK, T.; KUX, H. J. H. Urban land cover and land use classification of an informal settlement area using the open-source knowledge-based system InterIMAGE. Health, Risk & Society, v. 55, n. 1, p. 23-41, 2010. Disponível em: https://www.tandfonline.com/doi/abs/10.1080/14498596.2010.487640. Acesso em: 22 mai. 2021.
ORFEO TOOLBOX (OTB). (2021). Welcome to Orfeo ToolBox! Disponível em: https://www.orfeo-toolbox.org/CookBook/index.html. Acesso em: 12 jun. 2021.
PEREIRA, T. G. Aplicação de algoritmo computacional baseado em distância para classificações automáticas de tipologias florestais e de classes de volume de biomassa florestal nas Glebas Mamuru-Arapiuns localizadas no estado do Pará. 2016. 56 f. Dissertação (Mestrado) - Curso de Pós-graduação em Ciências Ambientais, Centro de Ciências Naturais e Tecnologia, Universidade do Estado do Pará, Belém, 2016. Disponível em: https://paginas.uepa.br/pcambientais/dissertacao_thamiris_pereira_turma_2014.pdf. Acesso em: 19 jun. 2021.
PINHO, C. M. D. Análise orientada a objetos de imagens de satélite de alta resolução espacial aplicada à classificação de cobertura do solo no espaço intra-urbano: o caso de São José dos Campos – SP. 2005. 179 p. Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2005.
PINHO, C. M. D.; UMMUS, M. E.; NOVACK, T. Extração de feições urbanas em imagens de alta resolução espacial a partir do estudo do comportamento espectral dos alvos. Revista Brasileira de Cartografia, v. 63, p. 619-627, 2011. Disponível em: http://www.seer.ufu.br/index.php/revistabrasileiracartografia/article/view/49213. Acesso em: 19 jun. 2021.
QUILAN, J. R. (1992). C4.5: programs for machine learning. San Francisco: Morgan Kaufmann Publishers. 302 p. Disponível em: https://link.springer.com/article/10.1007/BF00993309#citeas. Acesso em: 19 jun. 2021.

REUß, F. Detection of favelas in Brazil using texture parameters and machine learning. 2017. 94 f. Dissertação (Mestrado em Geospatial Technologies), Graz University Of Technology, Graz, 2017. Disponível em: https://elib.dlr.de/115220/1/Reuss_MA.pdf. Acesso em: 01 abr. 2021.

SANTOS, B. D.; PINHO, C. M. D. (2017). Análise de imagens de alta resolução espacial baseada em objetos para classificação de assentamentos precários. In: Simpósio Brasileiro de Sensoriamento Remoto, 18. (SBSR), 2017, Santos. Anais... São José dos Campos: INPE. p. 6226-6232. Disponível em: http://urlib.net/rep/8JMKD3MGP6W34M/3PSMCGH. Acesso em: 19 jun. 2021.

SIDDIQI, N. (2005). Intelligent Credit Scoring: building and implementing better credit risk scorecards. Hoboken: Wiley. 464 p.

SILVA, Maria Elisa Siqueira. (2016). Regressão linear: Introdução a Estatística Aplicada a Climatologia. 01 mar. 2016, 01 jun. 2016. 56 p. Notas de Aula.

STRAUSS, David. The many faces of logistic regression. The American Statistician, v. 46, n. 4, p. 321-326, 1992.

Terraview. Version 2.3.2. (2021). INPE. Disponível em: http://www.dpi.inpe.br/terralib5/wiki/doku.php?id=wiki:downloads. Acesso em: 15 mai. 2021.

United Nations Habitat (UN-Habitat) (org.). (2016). World Cities Report 2016: urbanization and development - emerging futures. Nairobi: United Nations. Disponível em: https://unhabitat.org/world-cities-report. Acesso em: 21 mai. 2021.

WURM, M.; TAUBENBÖCK, H. Detecting social groups from space: assessment of remote sensing-based mapped morphological slums using income data. Remote Sensing Letters, v.9, n.1, p. 41-50, 2017. Disponível em: https://www.tandfonline.com/doi/full/10.1080/2150704X.2017.1384586. Acesso em: 22 mai. 2021.