

Exploring the Potential of Machine Learning for Automatic Slum Identification from VHR Imagery

Juan C. Duque ^{1,*} , Jorge E. Patino ¹ and Alejandro Betancourt ^{2,3}

¹ Research in Spatial Economics (RiSE-Group), Department of Economics, Universidad EAFIT, Carrera 49 No. 7 Sur-50, 050022 Medellin, Colombia; jpatinoq@eafit.edu.co

² Department of Engineering (DITEN), University of Genova, 16145 Genova, Italy; a.betancourt@tue.nl

³ Department of Industrial Design, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

* Correspondence: jduquec1@eafit.edu.co; Tel.: +574-261-9354; Fax: +574-261-9294

Academic Editor: Qi Wang

Received: 26 July 2017; Accepted: 25 August 2017; Published: 30 August 2017

Abstract: Slum identification in urban settlements is a crucial step in the process of formulation of pro-poor policies. However, the use of conventional methods for slum detection such as field surveys can be time-consuming and costly. This paper explores the possibility of implementing a low-cost standardized method for slum detection. We use spectral, texture and structural features extracted from very high spatial resolution imagery as input data and evaluate the capability of three machine learning algorithms (Logistic Regression, Support Vector Machine and Random Forest) to classify urban areas as slum or no-slum. Using data from Buenos Aires (Argentina), Medellin (Colombia) and Recife (Brazil), we found that Support Vector Machine with radial basis kernel delivers the best performance (with $F2$ -scores over 0.81). We also found that singularities within cities preclude the use of a unified classification model.

Keywords: remote sensing; slum detection; machine learning

1. Introduction

According to [1], slums are the most deprived and excluded form of informal settlements. Slums are characterized by poverty and agglomerations of inadequate housing and are often located in hazardous urban areas. In 2016, approximately one in eight individuals lived in a slum. Although there was a decrease from 39% to 30% of urban population living in slums between 2000 and 2014, the absolute number of people living in urban slums continues to grow and it is a critical factor for the persistence of poverty in the world [2]. Moreover, the urban population of the world's two poorest regions, South Asia and Sub-Saharan Africa, is expected to double over the next 20 years, which suggests that the slum dwellers in those regions will grow significantly [1].

There has been a significant increase in the number of studies regarding the usefulness of remote sensing imagery to measure socioeconomic variables [3–6]. This trend is partly due to the increasing availability of satellite platforms, advances in methods and the decreasing costs of these images [7,8]. Remote sensing imagery may become an alternative source of information in urban settings for which survey data are scarce. In addition, this imagery may complement socioeconomic data that have been obtained from socioeconomic surveys [3]. The use of remote sensing data to estimate socioeconomic variables is based in the premise that the physical appearance of a human settlement is a reflection of the society that created it and is also based on the assumption that individuals who live in urban areas with similar physical housing conditions have similar social and demographic characteristics [9,10].

Slum detection or slum mapping is one of the most recurrent applications in this field of study; scholars have published a minimum of 87 papers in scientific journals over the last 15 years [8].

These studies have demonstrated that the physical characteristics of slums are distinguishable from the physical characteristics of formal settlements by using remote sensing data [11–13]. This is an important area of study because numerous local governments do not fully acknowledge the existence of slums or informal settlements [1], which hinders the formulation of policies to benefit the poor citizens of cities [8].

Numerous methods that make use of remote sensing imagery can be used to identify slum areas. Object based image analysis (OBIA) was, until recently, the most used method; other methods include visual interpretation, texture/morphology analysis and machine learning, which is more accurate and is often combined with OBIA [8]. Machine learning (ML) approaches generally combine textural, spectral and structural features [12]. The Random Forest classifier (RF) is one of the most popular ML methods for slum extraction that uses very high spatial resolution (VHR) imagery [12,14]. Support Vector Machine (SVM) and Neural Networks (NN) are also used for slum identification [8]. However, most of these ML algorithms are implemented at the pixel level and have limited viability when working with VHR imagery, in contrast to OBIA [15]. Appropriate ML methods are generally determined by the intuition of the researcher.

According to [8], most published studies describe the use of remote sensing to map slums and relied on expensive commercial imagery with near-infrared (NIR) information [16] or three-dimensional data such as LIDAR [13]. Numerous small cities in developing countries do not have the funds to purchase full satellite imagery and often use RGB data for data extraction via interpretation [15,17]. Google Earth (GE) imagery may be the only available source of aerial imagery for small local governments because these images are free to the public [18,19]. In addition, Google Earth provides historical VHR imagery for many locations, which may be useful for spatio-temporal urban analysis. According to Google Earth terms of service [20], GE imagery can be used for non-commercial purposes, and its use is specifically allowed for research papers and other related documents.

The purpose of this study is threefold. First, we explore the possibility of detecting slums within cities by using very high spatial resolution (VHR) RGB GE imagery, image feature extraction and OBIA techniques, without ancillary data. Second, using identical input data, we compare the performance of different ML algorithms to identify slums and determine which algorithm provides the optimal results. Third, we seek to identify a low-cost standardized method to detect slums that is also flexible, easy to automate and may be used in other urban settings with scarce data. We use data for three Latin American cities with different physical and climate conditions and different urban layout characteristics: Buenos Aires (Argentina), Medellin (Colombia), and Recife (Brazil).

The structure of this paper is as follows: Section 2 describes the methodology including a description of the data and the three classification models that are utilized in this study. Section 3 provides the results and a discussion of the implemented approach. Section 4 presents the primary conclusions, suggestions for future research and policy-making implications for local governments and authorities.

2. Methods

Our goal is to design an algorithm that can automatically identify the areas of a city that possess the urban characteristics of a Slum. This problem can be defined as a binary classification problem for which the inputs are features that have been extracted from GE images and the output is a binary variable that assumes the value of 1 if a particular area of the city is a slum and 0 otherwise. Figure 1 summarizes the proposed approach for detecting slums. This process begins with collecting the input data for the administrative boundaries. Data are obtained from Open Street Maps (OSM) and GE images for two different time instances (upper portion of the figure). The second stage of the process (middle portion of the figure) includes calculating spectral, textural and structural variables (i.e., the image feature extraction) from the GE images. During this stage, the images are discretized by overlapping a regular grid; the outer border is defined by the OSM boundary. This procedure generates Spatial Datasets (one per year, per city) that are composed of regular polygons with their

corresponding spectral, textural and structural variables. Finally, the third stage (lower segment of the figure) includes a classification analysis. The data for the most recent year are used to train the classification models and identify the best-performing model for slum identification. The optimal model is then applied to images from prior years to identify urban changes in the most important areas of each city.

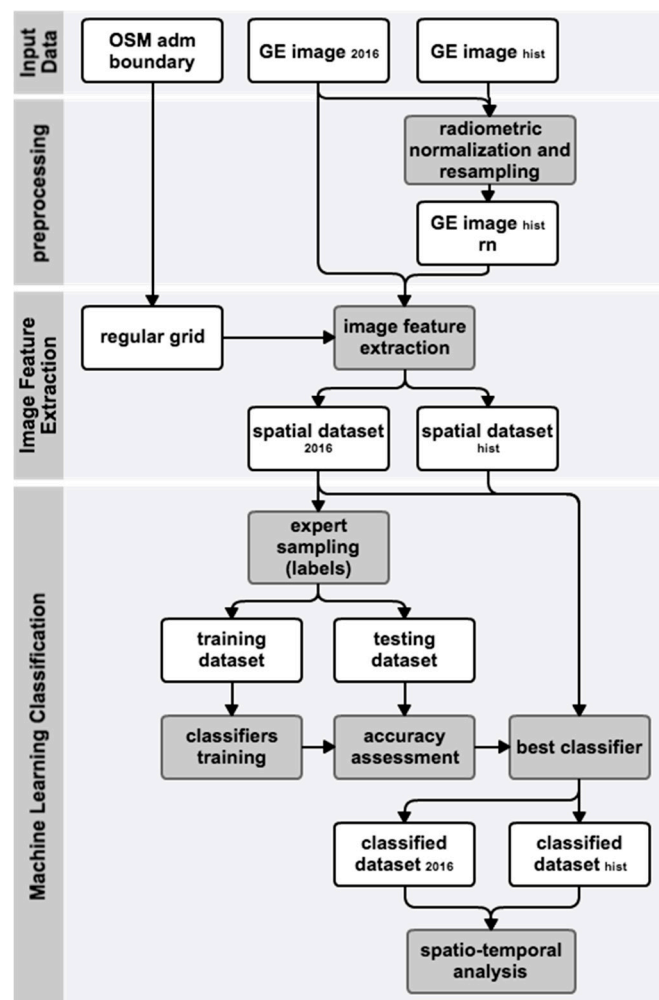


Figure 1. Flow diagram of the proposed approach for slum detection and change analysis.

2.1. The Data

We selected three Latin American cities to test the transferability of this approach: Buenos Aires, Argentina; Medellin, Colombia; and Recife, Brazil (Figure 2). These cities represent different climates, environmental conditions, and cultures and the use of different building materials. Buenos Aires is located at 34°35'59" S, 58°22'55" W at sea level and borders the La Plata river outlet to the ocean over plain lands and has a dry climate with marked seasons. Medellin is located at 6°14'41" N, 75°34'29" W in an intermountain valley at 1460 m above mean sea level and has a tropical, wet climate. Recife is located at 8°03'14" S, 34°52'51" W at sea level in a hilly terrain and has a tropical, wet climate. Table 1 provides general descriptions of these cities.



Figure 2. Cities included in the study.

Table 1. City descriptions.

Variable	Buenos Aires	Medellin	Recife
Population estimates for 2010 (people)	2,890,151	2,309,446	1,537,794
Area (km ²)	200	105	218
Density (people/km ²)	14,451	21,995	7040
Elevation (meters above sea level)	0	1480	0
Mean annual temperature (°C)	18	24	25
Average relative humidity (%)	72.0	68.3	79.8

Data sources: [21–23].

We downloaded the most updated (up to March 2016) GE images for each city and used a zoom level that was similar to VHR imagery with sub-meter pixel size. Google Earth imagery with very high spatial resolution is available for almost all urban areas worldwide. The VHR images were obtained from a number of providers or satellite platforms (e.g., Digital Globe, Geo Eye, and CNES/Astrium, among others). Images are captured by different sensors on different dates using different spatial resolutions; however, most of the images have a submeter pixel size and serve as natural-colored images that have three bands: red, green and blue (RGB). Because of the differences in platforms and different dates of acquisition, images captured at the same location on different dates will indicate differences through illumination conditions and color intensities. The GE images were georeferenced and rescaled between 0 and 255. We kept the preprocessing of the images to a minimum to gain speed in the workflow and to maintain the ease of automation of the whole approach.

Prior studies state that block-level spatial units of analysis are the most useful for urban planning purposes [13,24]. OpenStreetMap (OSM) data that layer streets and roads are useful to delineate urban blocks. However, in developing countries, cities' street networks are incomplete because of the high density and complexity of slum areas [13] or because areas that have been recently occupied have not been registered in all of the OSM datasets, as is the case for the northeastern section of Medellín city. In these instances, the delineation of urban blocks would add considerable processing time to the approach because it would require visual interpretation and manual digitalization of roads and pedestrian paths.

A simple alternative that can be automated is using a regular grid to detect slums from remote sensing imagery. Prior studies have used regular grids to extract, aggregate and classify image

data [8,25,26]. A regular grid in a vector, or fishnet, format can be drawn using any GIS software; the only necessary input is the boundary of the study area. This method could increase the speed of this study. We tested the use of two fishnets with different polygon sizes (a fishnet with square cells of 100 m on each side and another fishnet with square cells of 50 m on each side) for image feature extraction and classification. The results that were obtained with the 100 m grid outperformed the results that were obtained with the 50 m grid in regards to the correct classification of slum-like areas. The 100 m square cells are similar in size to actual urban blocks and have been recognized as an appropriate spatial unit of analysis to study intra-urban poverty for urban planning and policy making [24]. We downloaded the administrative boundaries of each city from OSM using QGIS [27] to define the extent of the study areas, and then created a regular grid of square cells with 100 m on each side over the urban areas of each city to extract the image features. The use of administrative boundaries to select the study areas could introduce bias in the identification of slums, as those areas located just outside the fringe will not be included in the analysis. As the focus of this work is to test the ability to identify slums from GE imagery in the three different Latin American cities using the same approach, rather than identify all the slum areas in a particular city, we used the administrative boundaries to select the areas in the same way for all three cases.

In addition, we selected well known slum areas in each city and downloaded cloud-free GE images for each sector from approximately a decade prior to test the approach's ability to analyze changes in slum areas. We attempted to capture images from the same city at two different points in time that were roughly a decade apart to determine if the proposed approach could identify changes that had taken place between the dates. This time span was restricted by the availability of Google Earth's VHR images for each city and by the quality of the available images, which can be affected by the presence of clouds and shadows. Historical VHR imagery provided by GE is also restricted to the availability of commercial VHR data, which was released after the launch of the Ikonos satellite in 1999. The most updated good quality VHR images available for Buenos Aires, Medellin, and Recife are from 2006, 2008, and 2008, respectively. Although images from other dates are available for these cities in GE, they were captured using medium spatial resolution platforms and are not suitable for extracting spatial pattern descriptors at the intra-urban scale.

The historical GE images were resampled to the identical pixel size of the 2016 images of each city, and we performed radiometric normalization between the historical images and the 2016 images; the 2016 images were used as a reference. Resampling and radiometric normalization were performed to obtain historical images with the identical pixel size and similar color intensity as the 2016 images (i.e., pixel values in each RGB band). Preprocessing the historical images simplifies the process to identify changes and differentiates between changes in intensity because of differences in illumination and atmospheric conditions.

2.1.1. Feature Extraction

Different image texture measures and spatial pattern descriptors (structure measures) have been used for differentiating slum areas from formal ones in several cities of developing countries around the world [3,12,24,28,29]. We used current GE images (obtained in March 2016) and the regular grid of each city to extract image information using FETEX 2.0. Figure 3 illustrates the outline of the urban areas for each city and selected sectors (500 by 500 m) illustrate the regular grid over the 2016 GE images. FETEX is an interactive software package that is used for image and object-oriented feature extraction [30] and is available on the Geo-Environmental Cartography and Remote Sensing Research Group website [31]. We calculated three sets of variables: a set of spectral features, a set of textural features and a set of structural features. The image features are extracted from the image by processing the pixels that are located within the same polygon without changing the image resolution or pixel values. Spectral features provide information regarding color; texture and structural features provide information regarding the spatial arrangement of the elements within the image. The urban layout of slum-like neighborhoods often displays a more organic, crowded and cluttered pattern than for more

formal and wealthy neighborhoods. Texture and structural features may help to differentiate between slum and no-slum areas [3,12,32,33].

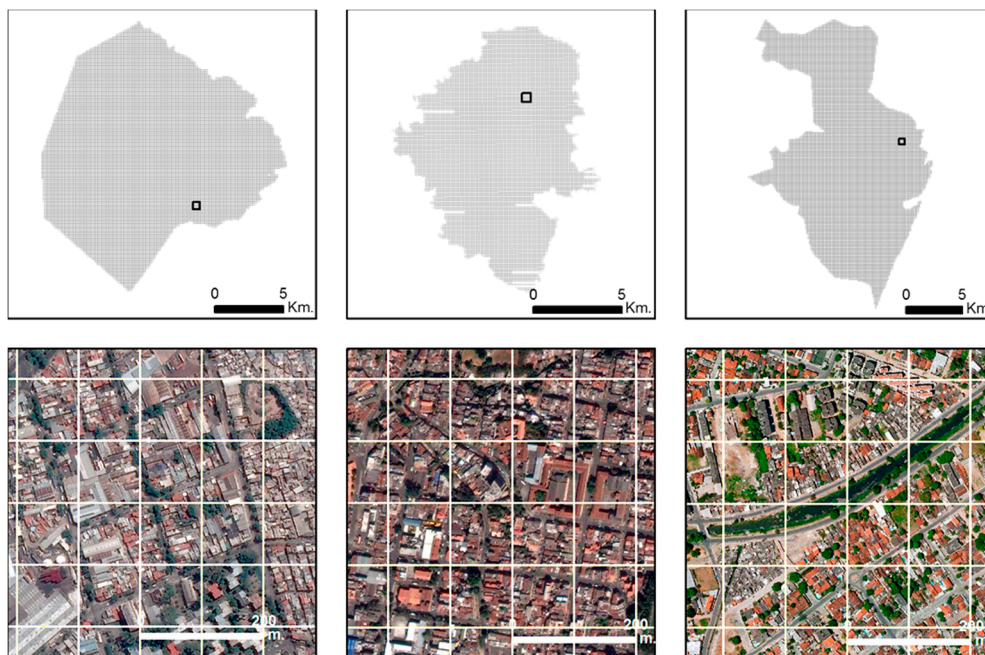


Figure 3. Urban areas and selected sectors showing the regular grid over the 2016 GE images of each city. From left to right: Buenos Aires, Medellin, and Recife.

Spectral features: Spectral features include the summary statistics of pixel values inside each polygon. These features provide information regarding the spectral response of objects, which differs for land coverage types, states of vegetation, soil composition, building materials, etc. [30]. We selected the mean and standard deviation for each RGB band and the majority statistic, to be extracted within this group. These features are easy to understand and provide better information about the spectral differences across the cities than other summary statistics (minimum, maximum, range, and sum).

Texture features: Textural features characterize the spatial distribution of intensity values of an image and provide information about contrast, uniformity, rugosity, etc. [30]. FETEX 2.0 performs texture feature extraction based on the Grey Level Co-occurrence Matrix (GLCM) and a histogram of pixel values inside each polygon. The kurtosis and skewness features are based on a histogram of the pixel values inside the polygon; the GLCM describes the co-occurrences of the pixel values that are separated at a distance of one pixel inside the polygon and is calculated considering the average value of four principal orientations, 0° , 45° , 90° and 135° , to avoid any effects of the orientation of the elements inside the polygon [30]. The GLCM of FETEX 2.0 was utilized to calculate a set of variables that were proposed by [34] and are widely used for image processing, including uniformity, entropy, contrast, inverse difference moment (IDM), covariance, variance, and correlation. The edgeness factor is another useful feature that represents the density of the edges of a neighborhood. The mean and standard deviation of the edgeness factor (MEAN EDG, and STDEV EDG) are also computed within this set of texture features in FETEX 2.0 [30].

Structural features: These features provide information regarding the spatial arrangement of elements inside the polygons in terms of the randomness or regularity of their distribution [30,35,36]. Structural features are calculated in FETEX using the experimental semivariogram approach. According to [30], the semivariogram quantifies the spatial associations of the values of a variable, measures the degree of spatial correlation between the different pixels of an image and is a suitable tool to determine regular patterns. FETEX 2.0 obtains the experimental semivariogram for each polygon by computing the mean of the semivariogram calculated in six different directions, from 0° to 150° .

in increments of 30°. Then, each semivariogram curve is smoothed using a Gaussian filter to reduce experimental fluctuations [30]. Structural features extracted from the semivariogram are based on the zonal analysis that is defined by a set of singular points on the semivariogram, such as the first maximum, the first minimum, and the second maximum [30]. For a full description of these features, see [30,35,36]. Table 2 provides a list of the remote sensing variables that are used for this analysis.

Table 2. Image-derived variables.

Group	Variable Name	Description
Spectral features	MEAN1	Mean of pixel values in band 1
	DEVST1	Standard deviation of pixel values in band 1
	MAJORITY1	Majority of pixel values in band 1
	MEAN2	Mean of pixel values in band 2
	DEVST2	Standard deviation of pixel values in band 2
	MAJORITY2	Majority of pixel values in band 2
	MEAN3	Mean of pixel values in band 3
	DEVST3	Standard deviation of pixel values in band 3
	MAJORITY3	Majority of pixel values in band 3
Texture features	MEAN_EDG	Mean of the edgeness factor
	DEVST_EDG	Standard deviation of the edgeness factor
	UNIFOR	GLCM uniformity
	ENTROP	GLCM entropy
	CONTRAS	GLCM contrast
	IDM	GLCM inverse difference moment
	COVAR	GLCM covariance
	VARIAN	GLCM variance
	CORRELAC	GLCM correlation
	SKEWNESS	Skewness value of the histogram
	KURTOSIS	Kurtosis value of the histogram
Structure features	RVF	Ratio variance at first lag
	RSF	Ratio between semivariance values at second and first lag
	FDO	First derivative near the origin
	SDT	Second derivative at third lag
	MFM	Mean of the semivariogram values up to the first maximum
	VFM	Variance of the semivariogram values up to the first maximum
	DMF	Difference between the mean of the semivariogram values up to the first maximum and the semivariance at first lag
	RMM	Ratio between the semivariance at first local maximum and the mean semivariogram values up to this maximum
	SDF	Second order difference between first lag and first maximum
	AFM	Area between the semivariogram value in the first lag and the semivariogram function until the first maximum

2.1.2. The Dataset

After the image features are extracted, the next step is to create the dataset. This process includes selecting a ground truth sample for each city. Each polygon of the sample is manually labeled as one of two categories: slum or no-slum. Ancillary information and prior studies were used as reference to identify slum areas in each city for sampling. A slum area can be considered a homogeneous zone with specific characteristics, but it can exhibit different appearances depending on the context [37]. However, most slum definitions relate to physical aspects of the built environment, which makes them comparable across settings. Although each city could have its own definition of slum, as pointed out by Taubenböck and Kraff, “the term slum is difficult to define, but if we see one, we know it” ([13], p. 15). The location of slum areas in Buenos Aires were identified on the “Caminos de la Villa” website [38] that provides an interactive map of the city and the location of recognized “villas” (slums). For Medellin we used the delineation of urban slums from [3,39], which is based on survey data and

the UN-Habitat global definition of slum [40]. The benchmark slum areas in Recife were identified using the work of [41], which shows the delimitation of widely recognized slum areas or “favelas” in that city. We visually checked the selected slum areas in each case to ensure that we were picking similar slum-like areas in all three cities. We then labeled as “slum” all the 100 m cells that overlapped with the slum areas from those already identified in the benchmarks. The sampling of no-slum areas in each city included different formal urban layouts such as high and low rise residential areas, parks, urban forests, green spaces, and commercial and industrial areas such as malls, transport facilities and factories. This binary classification scheme is common practice in remote sensing object-oriented approaches for identifying slum areas [13,28,29]. When benchmark information of slum areas is not available to construct a ground truth sample, practitioners must find reference information from local authorities or use an experienced interpreter who can visually determine slum and no-slum areas. Figure 4 provides the sampling spatial distribution for each city.

The final step in this stage is to divide the dataset into two sets: the training set that includes 60% of the sampled polygons for training and tuning the classification models and the testing set that includes 40% of the sampled polygons to evaluate the predictive capability of the classification models. Table 3 summarizes the composition of the datasets.

After the ground truth sampling was complete, we used the Kolmogorov–Smirnov (KS) test [42] and implemented the R package “kolmin” [43,44] to better understand the discriminating ability of the image-derived variables to differentiate slum areas from no-slum areas in each city.

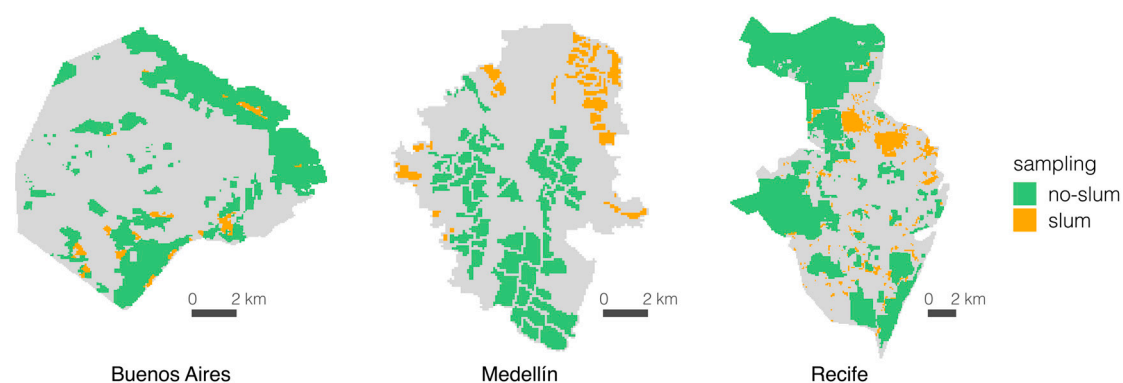


Figure 4. Sampling scheme of slum and no-slum areas for each city. From left to right: Buenos Aires, Medellín, Recife.

Table 3. Composition of the dataset (number of fishnet cells).

City	Total	Labeled	No Labeled	L. Slum	L. No Slum	Training	Testing
Buenos Aires	21,516	6558	14,958	369	6189	3934	2624
Medellin	10,255	2891	7364	602	2289	1734	1157
Recife	22,037	11,218	10,819	1274	9944	6730	4488

2.2. Classification Model

Classification literature is broad and multiple methods and algorithms have been proposed over recent decades [45]. In general, the primary goal is to develop a quantitative classification method that is capable of determining and generalizing the relationships between a set of variables (X) and a categorical variable (Y). For our specific classification problem, X is a matrix that includes the spectral, textural, and structural values of each polygon in the grid and Y is a categorical variable that assumes the value of either 1 or -1 if a polygon is a slum or no-slum, respectively. The capability of the classification method is determined by two factors: (i) the theoretical definition of the classification boundary of the classifier (e.g., linear and nonlinear); and (ii) the complexity of the data.

Based on the classification boundary, classifiers are commonly designated as either linear or nonlinear. Linear classifiers, such as logistic regression and linear SVMs, assume that the categorical variable (Y) can be obtained by exploiting a linear combination of the input features (X). Nonlinear classifiers generalize the boundary by adjusting polynomial boundaries, Gaussian kernels, or algorithmic criteria based on feature thresholding. Figure 5 illustrates a linear and nonlinear decision boundary. Nonlinear classifiers can capture more complex patterns from the data, but as a consequence, are more computationally complex than their linear counterparts and may be able to memorize the training data (overfitting).

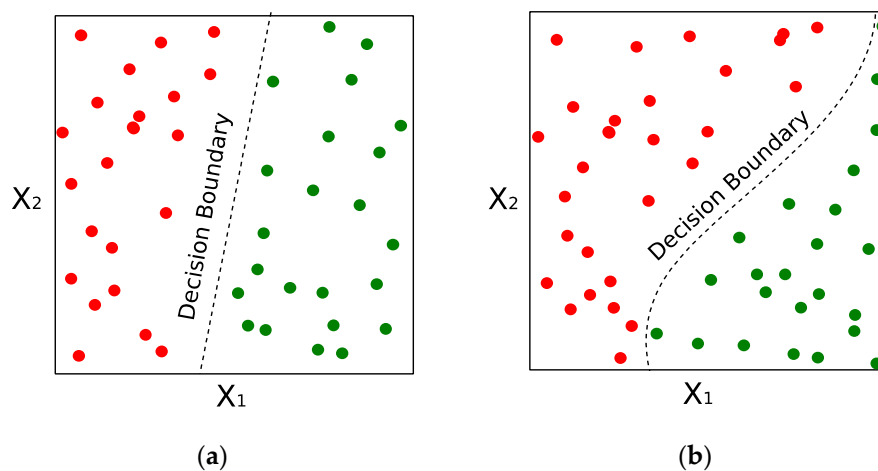


Figure 5. Linear and nonlinear classification boundary in 2D: (a) Linear Boundary on separable data; and (b) Nonlinear Boundary on separable data.

The intrinsic complexity of the data cannot be easily understood or described, particularly for high dimensional datasets. The most intuitive method to understand the data complexity is by visualizing its features and the respective classes. This approach is generally restricted to low dimensional data (2D or 3D) or simplified versions of the feature space that are obtained using manifold algorithms such as Principal Components Analysis (PCA), IsoMaps, or Self Organizing Maps [46]. A common approach that is used when working with dimensional data is to determine its complexity by comparing the capabilities of different classification algorithms to capture known patterns. To clarify, a simple classifier (linear) will perform poorly when using complex data (nonlinear) and complex classifiers (nonlinear) are able to use more complex data but have a large risk of overfitting. This risk is referred to as the bias–variance tradeoff [47], and it is commonly faced by adding a theoretical strategy known as regularization. The regularization strategy depends on the classification method and goes from the inclusion of additional terms in the error functions (e.g., Logistic Regression, SVM) to random disturbances in the training step and/or training data (e.g., Deep Neural Networks). In regards to the size of the training sets, there is no definitive number of observations that are required to train the models; this issue is commonly noted as a consequence of the complexity of the problem that is to be solved. Recent advances in data-science and deep-learning frequently refer to the benefits of large datasets; however, when data collection is expensive and time-consuming, a common practice is to observe changes in the evaluation criteria and sequentially increase the number of observations that are used to train the models. If the evaluation criteria do not improve (converge) as the number of training samples increases, then it is not necessary to collect additional training data.

Because our data have high dimensionality (30 features extracted per polygon) with unknown distributions and include data for three different cities, we explored two approaches for training our model to identify slums: (i) train a unique classifier on an unified dataset (i.e., without differentiating the cities) and then evaluate if the resulting slums are reliable; and (ii) use a multi-model approach by training the classifier in each city. Given the geographic and cultural differences as well as in the appearance of slums in these cities, fitting one method for slum identification in all the cities is

a huge challenge. However, it is important to test its feasibility in the search of robust tools for rapid urban slum detection with good performance in different settings. We analyze the performance of linear (Logistic Regression, linear SVM) and nonlinear classifiers (Polynomial and Radial Basis Kernel SVMs and Random Forests), which are available in the Python library Scikit-learn by [48].

The Logistic Regression (LR) is the most common linear classifier and is frequently used by policy makers in the econometric literature. This classifier is a mathematical approach whose primary goal is to use the logistic function to estimate the probability of a categorical value, Y , given the input features, X . For this classifier we used the Ridge regularization (known as L2), which compared to the Lasso regularization (known as L1) is less computationally expensive, provides a unique combination of coefficients, and, in case of correlated features, shrinks the estimates of the parameters but not to 0 [49–51]. The Support Vector Machine (SVM) is a popular non-probabilistic classification algorithm and is commonly recognized for its capability to maximize the margins between a decision boundary and the observations belonging to the particular categories. SVM, as a logistic regression, relies on a mathematical formulation to express the classification task as an optimization problem. This algorithm is highly popular in machine learning literature because of its ability to use nonlinear boundaries (kernels) from the theoretical formulation and its explicit goal of locating the boundaries as far as possible of the training data. In the experiment section, we use the polynomial kernel (SVMk), with k ranging from 1 to 5, and the radial basis kernel (SVMrbk). See [45] for a complete overview of the optimization procedure and more detailed information regarding the kernel functions. The regularization, in the case of the SVM, is defined as a constant that can be tuned to reduce overfitting. Finally, the Random Forest (RF), contrary to the Linear Regression and the SVM, makes a decision based on a sequential set of thresholding rules on the input space. Theoretically, a RF is an ensemble method that is formed by multiple decision trees. The RF decision is the average of the individual decisions of its trees, each of which is trained on bootstrap subset taken from the complete training data [52]. A decision tree is an algorithmic strategy that sequentially divides a feature space to fit the output variable [53]. For the results section we use the Least Squared Error (LSE) as the optimization function, the maximum depth of the trees is set to 10, and each random forest includes 10 decision trees. The use of the average to obtain the final decision endows RF, and in general all the ensemble methods, with an intrinsic robustness to overfitting. This is frequently pointed out as one their most significant advantages in Machine Learning literature.

2.3. Model Performance Assessment

Our comparison of the classifiers is based on the β score (F_β), which is a numeric performance defined by Equation (1), where the precision and recall are defined by Equations (2) and (3), respectively. Generally, precision measures the reliability of the slums that are detected (the purity of the regions that are detected as slum areas) and recall measures how efficiently the classifier retrieves the areas that are defined as slum areas (the number of slums that are detected). The F_β score, precision, and recall are bounded between 0 and 1; 1 represents a perfect classifier. The value of β must be selected according to the problem to be solved and is generally set to 0.5, 1 or 2. A value of $\beta = 0.5$ gives a larger weight to the precision and a value of $\beta = 2$ prioritizes the recall. In the remaining sections of this paper, β is defined as 2 (i.e., $F_{\beta=2}$) to give more importance to recall. This implies that, when classifying areas as slum or no-slum, we prefer type I errors over type II errors to prevent the vulnerable population from being ignored in the consideration.

$$F_\beta = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}} \quad (1)$$

$$\text{precision} = \frac{\text{TruePositives}}{\text{TruePositives} + \text{FalsePositives}} \quad (2)$$

$$\text{recall} = \frac{\text{TruePositives}}{\text{TruePositives} + \text{FalseNegatives}} \quad (3)$$

Once we have defined the best performing approach (unified or multi-model) and the best classifier, the next step is to tune the regularization constant to avoid overfitting the data and fine-tune the decision threshold to obtain the final $F2$ -scores. The regularization constant is exhaustively tuned by evaluating the $F2$ -score that is obtained while changing the regularization constant. The regularization constant that results in the highest $F2$ -score is defined as the final choice. The decision threshold is the value for which the classifier decides whether a particular observation is classified as slum or no-slum. The decision threshold is selected by using the Receiver Operating Characteristic (ROC) curve, which is a visualization of the False-Positives rates (X-axis) and True-Positives rates (Y-axis) while changing the decision thresholds. The machine learning bibliography suggests that the threshold is defined as the closest point to the upper-left corner of the ROC curve. It is important to note that the decision thresholds of the logistic regression reported in Section 3 are not bounded between 0 and 1, which is equivalent to using the X-axis for the final decision.

To ensure the tuning process is fair (regularization constant, decision threshold), only observations are used in the training dataset, which is accomplished by using cross-validation $F2$ -scores. To obtain the cross-validation $F2$ -scores, the first step is to divide the training dataset into k equal sized parts. On a single iteration, a classifier (with a specific regularization constant and decision threshold) is trained on $k - 1$ parts and tested in the remaining part to keep the $F2$ -score. This process is repeated k times to ensure that each part is used once for testing. The final cross-validation $F2$ -score is the average of the obtained $F2$ -score for each iteration. Our parameter selection is based on 10-fold cross-validation.

2.4. Slum Changes in Time

As stated above, we downloaded historical GE images for specific sectors of each city from roughly a decade ago (period $t - 1$) to perform change analysis. We selected identified slum sectors of 1 km^2 in the recent GE images (2016), and downloaded historical images of those sectors, from one decade ago, using historical imagery functionality in Google Earth. We applied relative radiometric normalization between the $t - 1$ image and the most recent image in each city. This process minimizes the differences in image data due to changes in atmospheric conditions, solar illumination, and view angles between images acquired at different dates. We extracted image features using the same regular grid of square cells and used the classifier model trained with the 2016 image-extracted data (period t) to classify each cell within the sector as either slum or no-slum. Then, cell by cell, we compared the results of the two dates (t vs. $t - 1$) and assigned different colors to differentiate the areas that were classified as slum for both dates, areas that were classified as no-slum for both dates, areas that were classified as no-slum for the $t - 1$ date but were classified as slum for the t date, and areas that were classified as slum for the $t - 1$ date and no-slum for the t date.

Following this rationale, we tested if the proposed approach could be useful to analyze slum dynamics over time by detecting areas that became slum areas, stable areas (no change), and areas that were slum areas and became no-slum areas by upgrading or through urban renovation processes.

3. Results and Discussion

3.1. Discriminating Image Features

The results of the Kolmogorov–Smirnov test indicate that the distributions of all image-derived variables are significantly different for the slum areas when compared to the no-slum areas. Figure 6 provides the boxplots for the five most discriminant image-extracted variables for each city (the results for the other variables are available upon request). It is notable that two of these five most discriminating variables were present for all three cities (SDF and CONTRAS) and all five variables are identical for Buenos Aires and Medellin (SDF, CONTRAS, IDM, MEAN EDG and FDO).

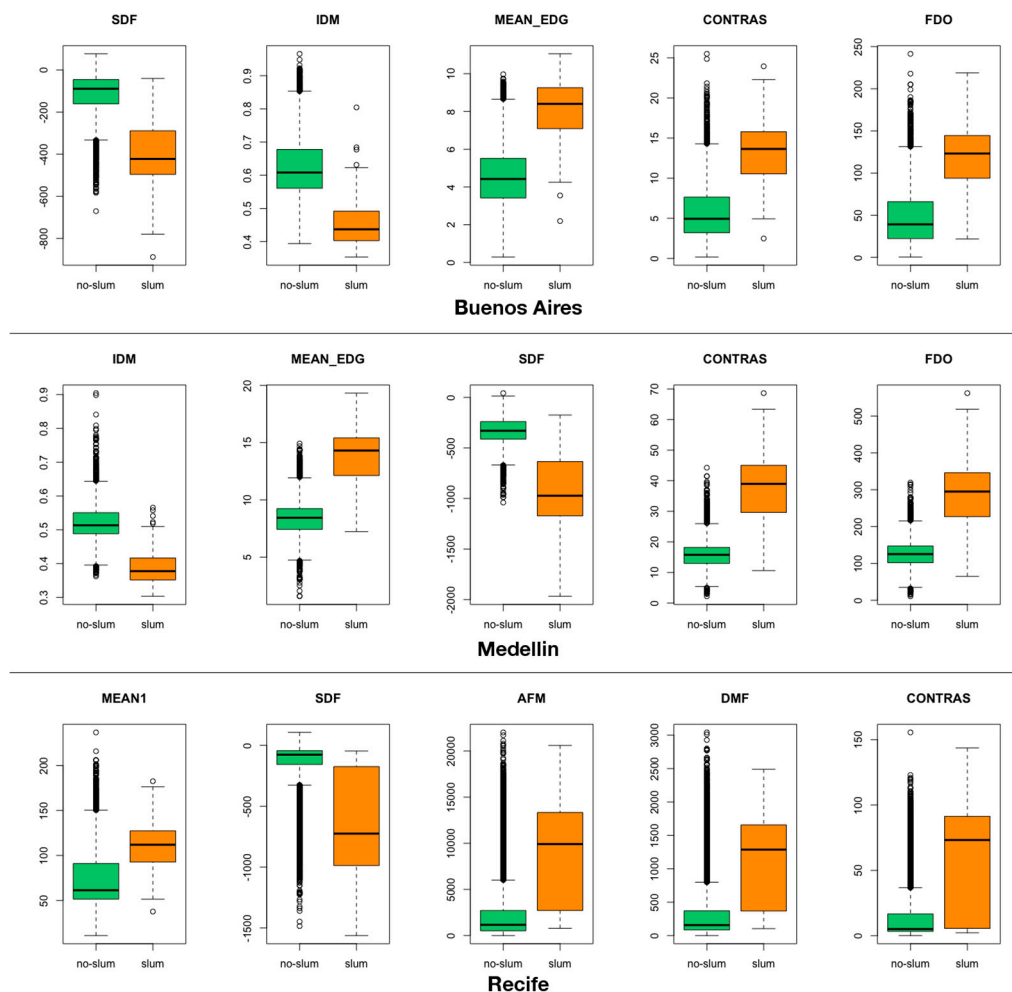


Figure 6. Boxplots of the distributions of the five most discriminant image-derived variables for each city. No-slum distributions are yellow and slum distributions are red. Variables are organized from left to right; the higher values of the Kolmogorov–Smirnov test are on the left and lower values are on the right.

These variables include textural and structural features, with the exception of MEAN1, which belongs to the spectral group and provides information about the mean of the intensity values in band 1, which corresponds to the red channel. SDF is a structure variable that provides information about homogeneity at short distances [35,36]. Slum areas demonstrate lower homogeneity than no-slum areas because they often include a variety of small dwelling units with different roof colors in close proximity to each other [3]. CONTRAS is a texture variable that provides information about the differences in color and intensity of the objects that are present in the image [34]. For the three cities that were included in this analysis, the slum areas had higher values for this variable than the no-slum areas. MEAN EDG is an aggregated measure of the density of edges present in an image [30]; the slum areas of these cities had higher values for this variable than no-slum areas because of the smaller sized dwelling units, narrower roads, and the presence of shadows between housing units and their surroundings. IDM is a texture measure that provides information about the general homogeneity [30,34]; slum areas are characterized by lower values of this feature than no-slum areas [3]. FDO is a structural feature that provides information about the variability of changes at short distances [35]; slum areas had higher values for this variable than no-slum areas because pixel values can change abruptly at short distances. AFM and DMF are structural features that are also related to the variability of the pixel values in the

image [35]; Recife slum areas had higher values for both of these features than no-slum areas, which implies that slum areas often display more variability and less homogeneity than no-slum areas.

3.2. Classification into Slum and No-Slum

For the first step in our experimental analysis, we use data for three cities to build a unified model. Table 4 provides the F_2 -score for each type of classifier in the testing set for each city. It is evident from the table that SVM_{rbk} is the best performing model. Regarding the polynomial SVMs (SVM_2 , ..., SVM_5), signals of underfitting occurred, particularly for the higher order models. The linear models (Logistic Regression and SVM_1) obtained a good classification score and did not show any signals of overfitting/underfitting. However, the Gaussian kernel of the SVM_{rbk} performs considerably better in all the cities. Finally, the poor performance of the Random Forest is noteworthy. The low classification scores provided by certain algorithms suggest the existence of singularities within cities that may complicate the identification of slums using a unified model. An additional boost in the performances can be obtained by carefully tuning each of the classifiers. For simplicity, this tuning is only applied to the best performing algorithms in the final part of this section.

Table 4. Unified Model for Slum Detection (F_2 -scores).

City	Log. R	SVM_1	SVM_2	SVM_3	SVM_4	SVM_5	SVM_{rbk}	RF
Buenos Aires	0.649	0.581	0.472	0.419	0.226	0.134	0.671	0.277
Medellin	0.757	0.516	0.794	0.817	0.821	0.841	0.872	0.550
Recife	0.671	0.552	0.559	0.565	0.564	0.592	0.803	0.516

Because of the differences in the cities' urban structures, we train a classification model for each city. Table 5 provides the testing F_2 -score for each model and each city. In this case, the best classification score is obtained by Logistic Regression and SVM_{rbk} ; both models achieved F_2 improvements between 2 and 5 points with respect to the unified model. The remainder of the models indicated certain improvements against their unified counterpart; however, their performance is still poor when compared to the Logistic Model and the SVM_{rbk} . These results confirm the intuition of structural differences in the features of the slums for each city that preclude the implementation of a unified model, which is in line with [33] who found morphological differences in spatial, spectral and textural characteristics of deprived areas in Mumbai. Figure 7 shows the distribution of the time required by each algorithm to classify a cell. As expected, the Logistic Regression is the fastest approach. The speed of the SVMs is comparable among them, even for those with complex kernels. Finally, the Random Forest is the slowest of the proposed algorithms. The results do not show significant differences between cities.

The next step is to remove signs of overfitting/underfitting of the best performing models and tune the decision threshold (th). This step only includes the Logistic Regression and the SVM_{rbk} . As explained in Section 2.2, the regularization term is selected by an exhaustive incremental search and the best threshold is selected by using the ROC curve. Table 6 provides the F_2 -scores of the default configuration (default), using only the tuned regularization term (Reg.) and using the tuned regularization term and the best threshold (Reg + th). The table confirms the benefits of the final tuning and allows us to conclude that the best strategy is to use a single model per city, include the regularization parameter and tune the decision threshold.

Table 5. Individual Model for Slum Detection (F_2 -scores).

City	Log. R	SVM_1	SVM_2	SVM_3	SVM_4	SVM_5	SVM_{rbk}	RF
Buenos Aires	0.737	0.642	0.649	0.678	0.652	0.642	0.688	0.596
Medellin	0.928	0.886	0.783	0.839	0.809	0.832	0.909	0.687
Recife	0.775	0.734	0.702	0.644	0.585	0.576	0.821	0.532

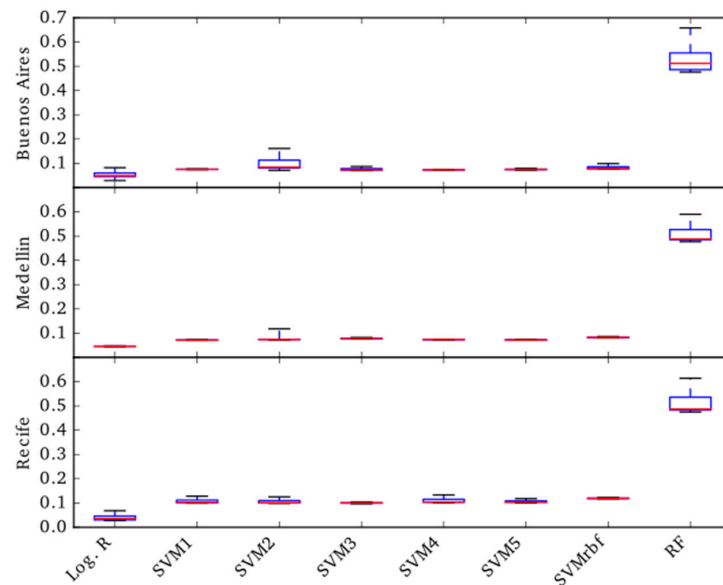


Figure 7. Time (s) taken by each method to classify a fishnet cell.

Table 6. Tuning the best classification models (F_2 -scores).

City	Log. Regression			SVM _{rbk}		
	Default	Reg.	Reg + th	Default	Reg.	Reg + th
Buenos Aires	0.737	0.759	0.715	0.688	0.813	0.817
Medellin	0.928	0.949	0.957	0.909	0.937	0.976
Recife	0.775	0.767	0.827	0.821	0.870	0.872

Figure 8 provides the F_2 -score for each SVM_{rbk} while changing the regularization term. The regularization value that maximizes the F_2 -score is set as the regularization term of the model. Figure 9 illustrates the ROC curves given the best regularization term. As previously explained, the decision threshold is set as the regularization term that generates the closest point to the upper-left corner of the curve. Using the data from Medellin, we found that the area below the ROC curve when using the 100 m fishnet was about 2% greater than the area below the ROC curve when using the 50 m fishnet (i.e., the ROC curve is slightly closer to the upper-left corner when using the 100 m fishnet). Finally, Table 7 provides the parameters that were selected for Table 6.

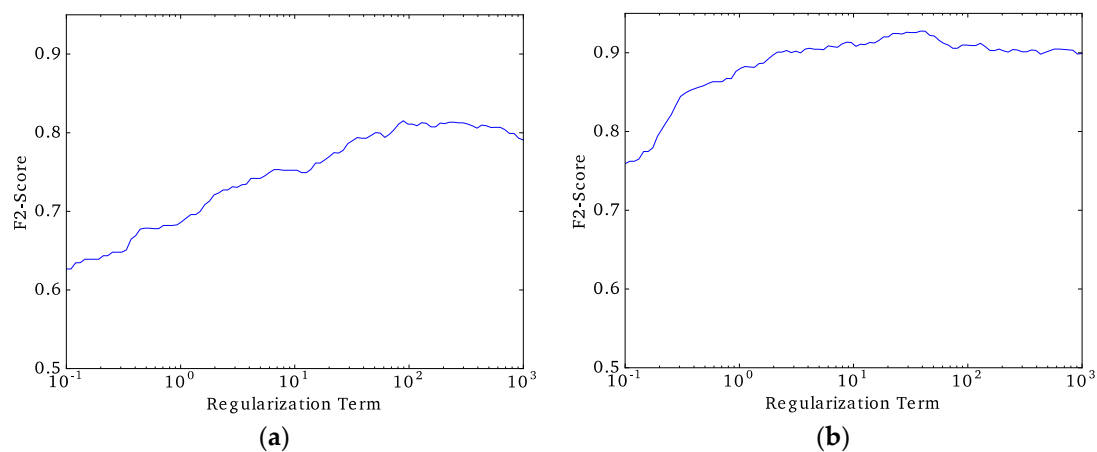


Figure 8. Cont.

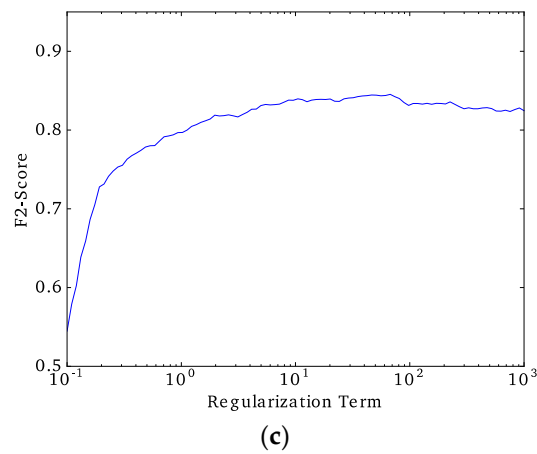


Figure 8. Tuning the regularization terms: (a) Buenos Aires; (b) Medellin; and (c) Recife.

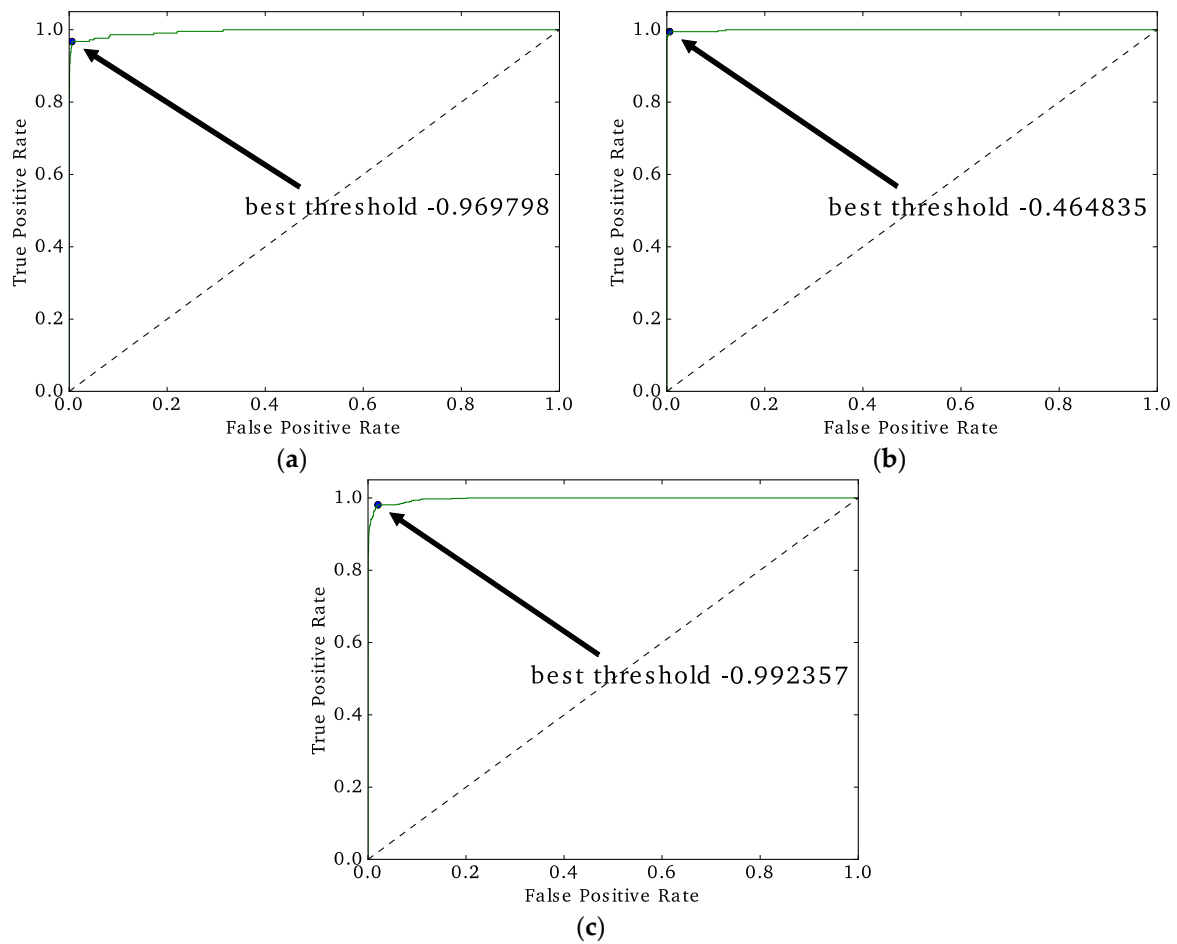


Figure 9. ROC curves used to select the best decision threshold: (a) Buenos Aires; (b) Medellin; (c) Recife. For visualization purposes, the x -axis is reported in logarithmic scale. The final regularization term is reported in Table 7.

Table 7. Final parameters of the logistic regression and the SVM_{rbk}. The values of the decision threshold for the logistic regression are reported in the x -axis, and are not bounded between 0 and 1. For both cases (Logistic Regression and SVM), more negative thresholds indicate the classifier is more prone to set an observation as slum.

City	Log. Regression		SVM _{rbk}	
	Reg. Term	Threshold	Reg. Term	Threshold
Buenos Aires	12.618	−3.111	394.42	−0.969
Medellin	20.092	−1.494	16.681	−0.464
Recife	225.70	−1.788	18.307	−0.992

Figure 10 provides the maps of detected slum areas using the classification process of the 2016 GE images for each city. The percentage of urban areas that are covered by slums is 24% for Buenos Aires and 36% for both Medellin and Recife. The spatial patterns changes in slums areas across the cities is as follows. In Buenos Aires, the slums are dispersed in little pockets throughout the territory. The slums emerge in intra-urban vacant lots and even in the periphery of industrial areas. In Medellin, slums are located in peripheral green areas adjacent to existing slums. This pattern is one of the consequences of armed conflict with guerrillas in rural areas; large groups of individuals were forced to move to the periphery of major cities in Colombia. In Recife, the slums are distributed in large clusters throughout the city and near highways (e.g., Highway BR-101) and rivers (the Capibaribe River).

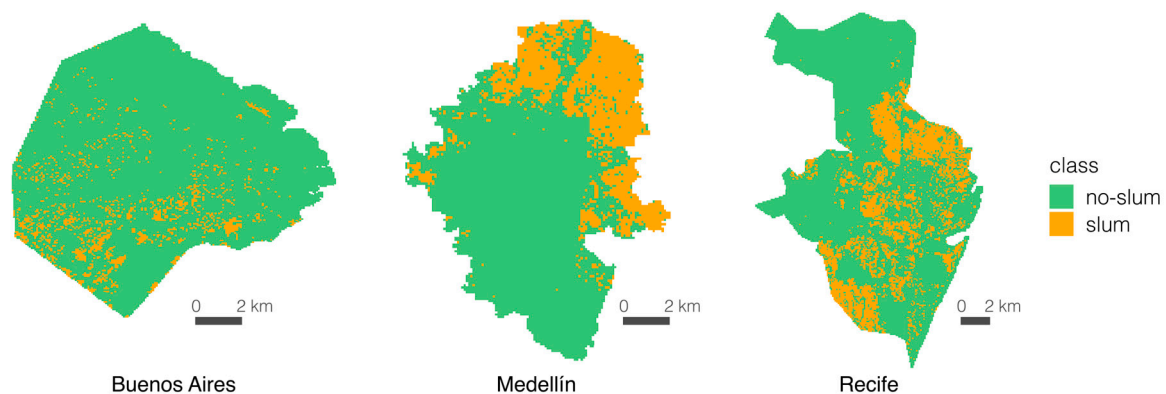


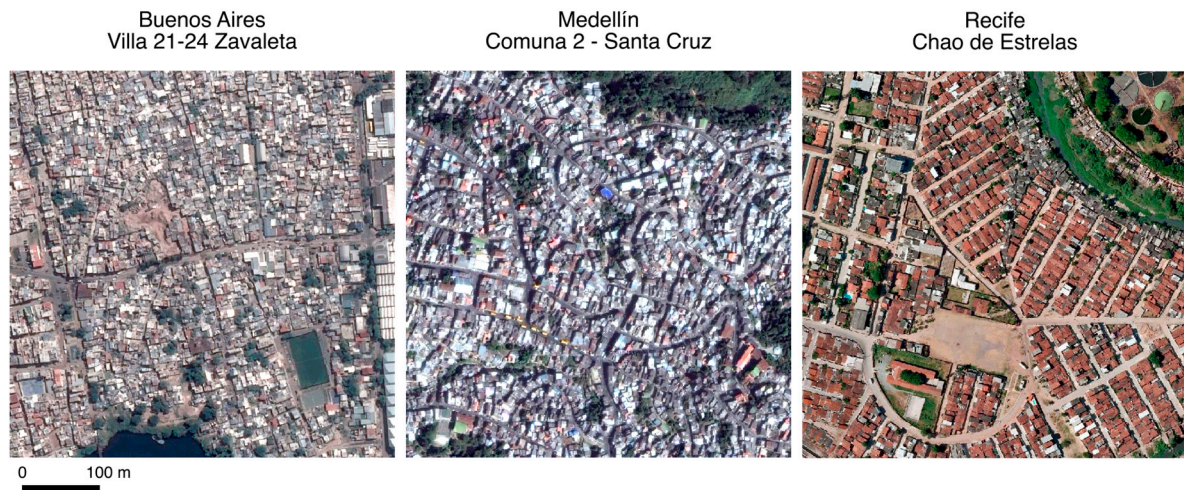
Figure 10. Classification results for 2016 GE images of each city to slum and no-slum areas.

3.3. Accuracy Assessment of Classification Results

The confusion matrix of the classification results presented in Table 8 illustrates the magnitude of the overestimation (no-slum areas classified as slum areas) and underestimation (slum areas classified as no-slum areas) for the testing dataset for each city. Buenos Aires resulted in less than 3% of overestimation and approximately 1% of underestimation; Medellin was the best case, with 2% overestimation and no underestimation; and Recife indicated 4.5% overestimation and less than 1% underestimation. Figure 11 provides the known slum sectors for each city using an identical spatial scale: Villa Zavaleta (21–14) in Buenos Aires, Comuna Santa Cruz in Medellin, and Chao de Estrelas in Recife. Table 9 provides a general characterization of slum areas for each city in terms of the image-derived features to better understand the classification results. The slums in all three cities are composed of clusters of small dwelling units and very few vegetated areas.

Table 8. Confusion Matrix of the SVM_{rbk} with the parameters reported in Table 7.

	(a) Buenos Aires		(b) Medellin		(c) Recife	
	Predicted No-Slum	Predicted Slum	Predicted No-Slum	Predicted Slum	Predicted No-Slum	Predicted Slum
No-Slum	2387	84	903	27	3800	203
Slum	17	136	0	227	32	453

**Figure 11.** Sectors of known slum areas in each city. From left to right: Buenos Aires, Medellin and Recife.**Table 9.** General characteristics of slums in the analyzed cities.

Observable Feature	Buenos Aires	Medellin	Recife
Short distance heterogeneity	high	High	medium to high
Large distance homogeneity	very high	very high	medium to high
Roofing material diversity	high	High	medium to high
Vegetation coverage	very low	Low	low
Urban layout pattern	organic	very organic	regular
Crowdedness	very high	High	medium to high
Dwelling size	very small (<60 m ²)	very small (<60 m ²)	small (<80 m ²)
Roads material	not paved	Paved	not paved
Roads width	very narrow (<6 m)	narrow (5–10 m)	narrow (5–10 m)

However, as expected from the boxplots of Figure 6, the slums in Buenos Aires and Medellin are similar when compared to slums in Recife. The slum areas in Buenos Aires and Medellin are characterized by high heterogeneity at short distances, high homogeneity at large distances and similar organic patterns. This means that there are different objects in close proximity (centimeters or few meters), but the same pattern is observed at larger distances across the territory (tens of meters); e.g., settlements that are made up of small dwellings with different building and roofing materials, which are located very close to each other, and with the same general pattern over the settlement or neighborhood. However, the slums in Buenos Aires are more cluttered than in Medellin. The slums in Recife have more homogeneity in color because most of the roofs are made of clay tiles or similar products, which explains the high discriminating power of the variable MEAN1 for this city because band 1 records the intensity values of the red channel in the visible spectrum and slum areas have many pixels with the same red tone in this city. Furthermore, the slums in Recife have more regularity in the spatial pattern of the urban layout than the slums in Buenos Aires and Medellin.

The lower score obtained for Buenos Aires may be explained by the quality of the input GE image and because the no-slum areas of the city had similar characteristics to the slum areas. The Buenos

Aires' GE image indicates low contrast (differences in color intensity tend to be low across the image), which could lower the quantifiable differences between the slum and no-slum areas. Figure 12 features a Zavaleta villa next to a no-slum area. Both areas indicate that very few vegetated areas exist between buildings and high heterogeneity occurs at short distances, but the no-slum sector has more regularity in the spatial pattern of the urban layout and large homogeneous surfaces are interspersed with clusters of smaller buildings.

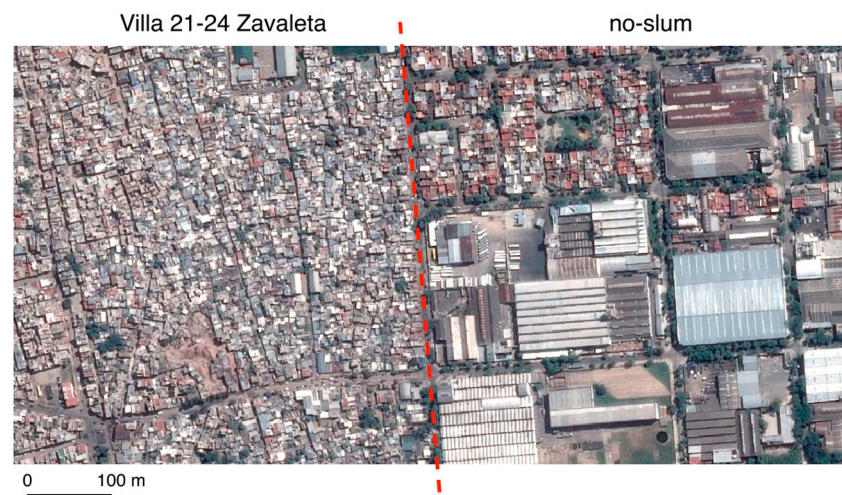


Figure 12. Slum sector in Buenos Aires (Villa Zavaleta) compared to the adjacent no-slum area. The red line indicates the slum boundary as mapped in [38].

3.4. Temporal Analysis

Figure 13 provides the results of the temporal analysis of the selected sectors of one square kilometer for each city. This approach is useful to provide information from a global perspective regarding areas that have changed from no-slum to slum and vice versa between the analyzed dates. However, as in the implemented algorithm, recall was given priority over the precision for a good identification of the more problematic regions within the city; we expected to obtain false positives in the classification results. These false positives adversely impact the interpretation at a detailed scale or on a cell-by-cell (of the regular grid) basis and can mask changes that an interpreter might see when comparing the two GE images for each sector. Because of the lack of reference data, we assessed the classification results of historic images by using on-screen interpretation. The obtained overall accuracies were 92% for Buenos Aires, 90% for Medellin and 72% for Recife. However, the results indicate certain interesting general trends: In Buenos Aires, slum areas tends to grow by using available space that is adjacent to already existing slum areas (e.g., vacant spaces between existing structures, zones adjacent to railroad tracks, and even parking lots in industrial areas). In Medellin, the slum areas grow by occupying undeveloped land on the edge of the urban perimeter. In 2008, we note the first part of the slum and, in 2016, the slum areas extend southwards over the adjacent “green” or “free” areas. Finally, in Recife, certain slum areas disappeared between 2008 and 2016; certain slum areas were located adjacent to the river and were removed to allow for a road on the riverbank. Certain other areas, with green or bare soil, were occupied by either slum areas or formal developments.

The proposed approach is optimal for identifying recently informally occupied urban areas that have slum characteristics versus changes due to slum upgrading processes. When slum areas are upgraded, this process often includes improving dwelling units and offering public services [2]; this process less often includes the modification of an urban layout because that implies relocation of a population and many slum residents fear that redevelopment will leave them homeless [54]. In this regard, upgrading processes that do not significantly change the spatial pattern and texture of

the urban areas cannot be determined by this approach because most of the image-derived features quantify aspects of the urban scene that are related to the spatial pattern and texture of the urban layout.

This workflow worked well for slum detection for a single date, but it did not work as well for spatio-temporal analysis. Although the historical images were resampled to match the pixel size of the updated reference images and they were normalized to match the color intensity, differences still exist for view angles, lighting, and vegetation phenology cycles between images that can affect vegetation appearance and shadow extent and affect the values of the image-derived features and the classification results. To minimize the differences in view angles and vegetation phenology, a practitioner must use historical images captured on the same day as the reference or an updated image; however, this is nearly impossible to control using the data available from Google Earth. Commercial satellite VHR imagery is more appropriate for this purpose because it can be acquired for specific dates and can match the day of the original image and minimize differences. The use of transfer learning methods recently introduced to remote sensing classification problems could overcome these issues [55].

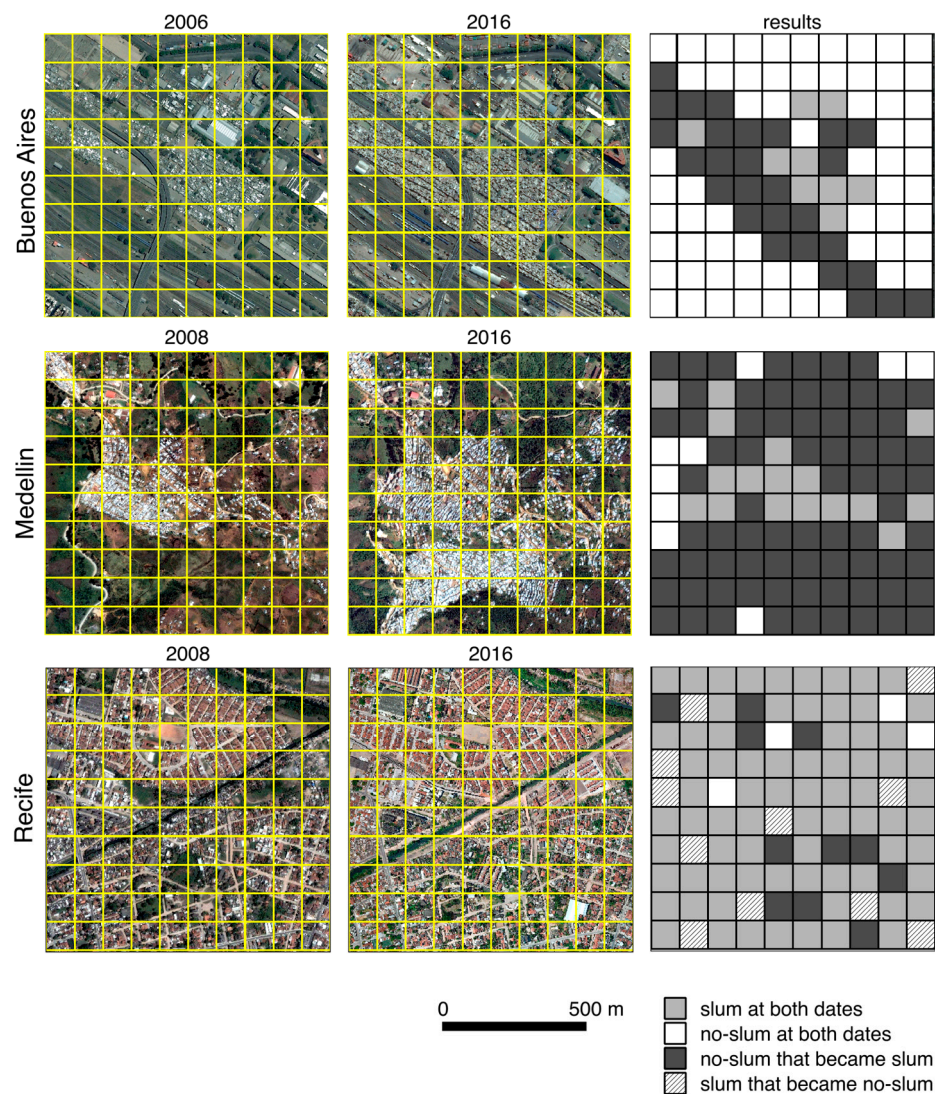


Figure 13. Classification results for historical GE images of selected sectors for each city.

4. Conclusions

This study explored implementing a low-cost standardized method for slum detection using spectral, texture and structural features extracted from VHR GE imagery that was utilized as input data and assessed the capability of three ML algorithms to classify urban areas as either slum or no-slum. Using data from Buenos Aires (Argentina), Medellin (Colombia), and Recife (Brazil), we determined that Support Vector Machine with radial basis kernel (SVM_{rbk}) performed the best with a *F2*-score over 0.81.

In addition, we determined that the specific characteristics of each city are important to consider and preclude the use of a unified classification model. The ML algorithms performed best for Medellin and Recife and resulted in *F2*-scores of 0.98 and 0.87, respectively. The image-derived features performed better for slum detection in these cities because their slum areas have a different spatial pattern and texture than no-slum areas and exhibit significant variations in the use of building and roofing materials.

The proposed workflow requires more sophistication to properly track changes over time because for the implemented ML algorithms, recall was given a higher priority than precision to obtain a good identification of the more problematic regions within the cities; false positives occurred in the classification results that adversely impact the change analysis between different dates. However, the proposed approach did identify recently and informally occupied urban areas that possessed slum characteristics, where the changes in local heterogeneity and the spatial pattern are clearly identified and were different from occupied formal areas. Changes in the slum status of an area because of upgrading processes would still be difficult to identify because those processes do not significantly change the spatial pattern and texture of the urban areas, which are the aspects quantified by the image-derived variables.

A suggestion for future studies is to use algorithms for object and scene recognition on images that are obtained from Google Street View to generate a new set of features that can improve the performance of our classification models. Street views and satellite imagery for slum identification can also be an important tool for supporting programs such as the Trust Fund for the Improvement of Family Housing that is led by the Development Bank of Latin America and the Foundation in Favor of Social Housing.

Acknowledgments: We gratefully acknowledge the financial support from CAF—Development Bank of Latin America. We also thank to the Google Company for providing freely images for research purpose. Finally, we thank the anonymous *Remote Sensing* reviewers for their insightful and helpful comments during the review process. The usual disclaimer applies.

Author Contributions: Juan C. Duque, who was the PI of the project that funded this research, structured the paper and, with Jorge Patino wrote the majority of the paper. Jorge Patino was in charge of the image processing and Alejandro Betancourt ran the machine learning algorithms.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. UN-Habitat. Habitat III Issue Papers 22—Informal Settlements. In Proceedings of the United Nations Conference on Housing and Sustainable Urban Development, N'Djamena, Chad, 8 May 2015.
2. UN-Habitat Slum Almanac 2015/2016. *Tracking Improvement in the Lives of Slum Dwellers*; Technical Report; Participatory Slum Upgrading Programme: Nairobi, Kenya, 2016.
3. Duque, J.C.; Patino, J.E.; Ruiz, L.A.; Pardo-Pascual, J.E. Measuring intra-urban poverty using land cover and texture metrics derived from remote sensing data. *Landsc. Urban Plan.* **2015**, *135*, 11–21. [[CrossRef](#)]
4. Engstrom, R.; Sandborn, A.; Yu, Q.; Graesser, J. Assessing the relationship between spatial features derived from high resolution satellite imagery and census variables in Accra, Ghana. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2015), Milan, Italy, 26–31 July 2015; pp. 2544–2547.

5. Sandborn, A.; Engstrom, R.N. Determining the Relationship between Census Data and Spatial Features Derived from High-Resolution Imagery in Accra, Ghana. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 1970–1977. [[CrossRef](#)]
6. Engstrom, R.; Newhouse, D.; Haldavanekar, V.; Copenhaver, A.; Hersh, J. Evaluating the relationship between spatial and spectral features derived from high spatial resolution satellite data and urban poverty in Colombo, Sri Lanka. In Proceedings of the 2017 Joint Urban Remote Sensing Event (JURSE), Dubai, UAE, 6–8 March 2017; pp. 1–4.
7. Patino, J.E.; Duque, J.C. A review of regional science applications of satellite remote sensing in urban settings. *Comput. Environ. Urban Syst.* **2013**, *37*, 1–17. [[CrossRef](#)]
8. Kuffer, M.; Pfeffer, K.; Sliuzas, R. Slums from Space—15 Years of Slum Mapping Using Remote Sensing. *Remote Sens.* **2016**, *8*, 455. [[CrossRef](#)]
9. Jain, S. Remote sensing application for property tax evaluation. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *10*, 109–121. [[CrossRef](#)]
10. Taubenböck, H.; Wegmann, M.; Roth, A.; Mehl, H.; Dech, S. Urbanization in India—Spatiotemporal analysis using remote sensing data. *Comput. Environ. Urban Syst.* **2009**, *33*, 179–188. [[CrossRef](#)]
11. Kuffer, M.; Barros, J. Urban morphology of unplanned settlements: The use of spatial metrics in VHR remotely sensed images. *Procedia Environ. Sci.* **2011**, *7*, 152–157. [[CrossRef](#)]
12. Kuffer, M.; Pfeffer, K.; Sliuzas, R.; Baud, I. Extraction of Slum Areas from VHR Imagery Using GLCM Variance. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 1830–1840. [[CrossRef](#)]
13. Taubenböck, H.; Kraff, N.J. The physical face of slums: A structural comparison of slums in Mumbai, India, based on remotely sensed data. *J. Hous. Built Environ.* **2014**, *29*, 15–38. [[CrossRef](#)]
14. Wurm, M.; Weigand, M.; Schmitt, A.; Geiß, C.; Taubenböck, H. Exploitation of textural and morphological image features in Sentinel-2A data for slum mapping. In Proceedings of the 2017 Joint Urban Remote Sensing Event (JURSE), Dubai, UAE, 6–8 March 2017.
15. Kohli, D.; Sliuzas, R.; Stein, A. Urban slum detection using texture and spatial metrics derived from satellite imagery. *J. Spat. Sci.* **2016**, *61*, 405–426. [[CrossRef](#)]
16. Owen, K.K.; Wong, D.W. An approach to differentiate informal settlements using spectral, texture, geomorphology and road accessibility metrics. *Appl. Geogr.* **2013**, *38*, 107–118. [[CrossRef](#)]
17. Guo, Z.; Shao, X.; Xu, Y.; Miyazaki, H.; Ohira, W.; Shibasaki, R. Identification of village building via Google Earth images and supervised machine learning methods. *Remote Sens.* **2016**, *8*, 271. [[CrossRef](#)]
18. Yang, X.; Jiang, G.-M.; Luo, X.; Zheng, Z. Preliminary mapping of high-resolution rural population distribution based on imagery from Google Earth: A case study in the Lake Tai basin, eastern China. *Appl. Geogr.* **2012**, *32*, 221–227. [[CrossRef](#)]
19. Hu, Q.; Wu, W.; Xia, T.; Yu, Q.; Yang, P.; Li, Z.; Song, Q. Exploring the use of Google Earth imagery and object-based methods in land use/cover mapping. *Remote Sens.* **2013**, *5*, 6026–6042. [[CrossRef](#)]
20. Google Maps/Google Earth Additional Terms of Service. Available online: https://www.google.com/intl/ALL/help/terms_maps.html (accessed on 2 May 2017).
21. Ciudad de Buenos Aires. Available online: www.buenosaires.gob.ar/laciudad/ciudad (accessed on 12 May 2017).
22. Alcaldía de Medellín. Available online: www.medellin.gov.co (accessed on 14 May 2017).
23. Instituto Brasileiro de Geografia. Available online: <https://cidades.ibge.gov.br/> (accessed on 18 May 2017).
24. Cabrera-Barona, P.; Wei, C.; Hagenlocher, M. Multiscale evaluation of an urban deprivation index: Implications for quality of life and healthcare accessibility planning. *Appl. Geogr.* **2016**, *70*, 1–10. [[CrossRef](#)]
25. Schöpfer, E.; Tiede, D.; Lang, S.; Zeil, P. Damage assessment in townships using VHSR data the effect of Operation Murambatsvina/Restore Order in Harare, Zimbabwe. In Proceedings of the Urban Remote Sensing Joint Event (URS), Paris, France, 11–13 April 2007; pp. 2–6.
26. Tapiador, F.J.; Avelar, S.; Tavares-Corrêa, C.; Zah, R. Deriving fine-scale socioeconomic information of urban areas using very high-resolution satellite imagery. *Int. J. Remote Sens.* **2011**, *33*, 6437–6456. [[CrossRef](#)]
27. QGIS Geographic Information System. Available online: <http://www.qgis.org/en/site/forusers/visualchangelog28/index.html> (accessed on 2 July 2017).
28. Rhinane, H.; Hilali, A.; Berrada, A.; Hakdaoui, M. Detecting slums from SPOT data in Casablanca Morocco using an object based approach. *J. Geogr. Inf. Syst.* **2011**, *3*, 209–216. [[CrossRef](#)]

29. Kit, O.; Lüdeke, M. Automated detection of slum area change in Hyderabad, India using multitemporal satellite imagery. *ISPRS J. Photogramm. Remote Sens.* **2013**, *83*, 130–137. [[CrossRef](#)]
30. Ruiz, L.A.; Recio, J.A.; Fernández-Sarría, A.; Hermosilla, T. A feature extraction software tool for agricultural object-based image analysis. *Comput. Electron. Agric.* **2011**, *76*, 284–296. [[CrossRef](#)]
31. Geo-Environmental Cartography and Remote Sensing Group. Available online: <http://cgat.webs.upv.es> (accessed on 10 May 2017).
32. Kuffer, M.; Sliuzas, R.; Pfeffer, K.; Baud, I. The utility of the co-occurrence matrix to extract slum areas from VHR imagery. The case of Mumbai, India. In Proceedings of the Joint Urban Remote Sensing Event (JURSE), Lausanne, Switzerland, 30 March–1 April 2015; pp. 3–6.
33. Kuffer, M.; Pfeffer, K.; Sliuzas, R.; Baud, I.; Maarseveen, M.V. Capturing the Diversity of Deprived Areas with Image-Based Features: The Case of Mumbai. *Remote Sens.* **2017**, *9*, 384. [[CrossRef](#)]
34. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621.
35. Balaguer, A.; Ruiz, L.A.; Hermosilla, T.; Recio, J.A. Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification. *Comput. Geosci.* **2010**, *36*, 231–240. [[CrossRef](#)]
36. Balaguer-Beser, A.; Ruiz, L.A.; Hermosilla, T.; Recio, J.A. Using semivariogram indices to analyse heterogeneity in spatial patterns in remotely sensed images. *Comput. Geosci.* **2013**, *50*, 115–127. [[CrossRef](#)]
37. Kohli, D.; Stein, A.; Sliuzas, R. Uncertainty analysis for image interpretations of urban slums. *Comput. Environ. Urban Syst.* **2016**, *60*, 37–49. [[CrossRef](#)]
38. Caminos de la Villa. Available online: <https://www.caminosdelavilla.org> (accessed on 2 July 2017).
39. Duque, J.C.; Royuela, V.; Noreña, M. A stepwise procedure to determine a suitable scale for the spatial delineation of urban slums. In *Defining the Spatial Scale in Modern Regional Analysis, Advances in Spatial Science*; Fernandez, E., Rubiera Morollón, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 237–254.
40. UN-Habitat. *Analytical Perspective of Pro-Poor Slum Upgrading Frameworks*; United Nations Human Settlements Programme: Nairobi, Kenya, 2006.
41. Da Silva, R.D.C.O.; Shaw, K. Cartography of the Favela. Available online: www.adrmarketplace.com/Cidade/Cartography.pdf (accessed on 26 July 2017).
42. Marsaglia, G.; Tsang, W.W.; Wang, J. Evaluating Kolmogorov's Distribution. *J. Stat. Softw.* **2003**, *8*, 1–4. [[CrossRef](#)]
43. Carvalho, L. An Improved Evaluation of Kolmogorov's Distribution. *J. Stat. Softw.* **2015**, *65*, 1–8. [[CrossRef](#)]
44. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2013. Available online: <http://www.R-project.org/> (accessed on 21 November 2016).
45. Hastie, T.; Tibshirani, R.J.; Friedman, J. *The Elements of Statistical Learning*, 10th ed.; Springer: Berlin, Germany, 2009.
46. Betancourt, A.; Diaz-Rodriguez, N.; Barakova, E.; Marcenaro, L.; Rauterberg, M.; Regazzoni, C. Unsupervised Understanding of Location and Illumination Changes in Egocentric Videos. *arXiv* **2016**, arXiv:1603.09200.
47. Geman, S.; Bienenstock, E.; Doursat, R. Neural Networks and the Bias/Variance Dilemma. *Neural Comput.* **1992**, *4*, 1–58. [[CrossRef](#)]
48. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python Gaël Varoquaux. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
49. Cessie, S.L.; Houwelingen, J.V. Ridge estimators in logistic regression. *Appl. Stat.* **1992**, *41*, 191–201. [[CrossRef](#)]
50. Tibshirani, R. Regression Selection and Shrinkage via the Lasso. *J. R. Stat. Soc. Ser. B* **1996**, *58*, 267–288.
51. Ng, A.Y. Feature selection, L 1 vs. L 2 regularization, and rotational invariance. In Proceedings of the Twenty-First International Conference on Machine Learning, Banff, AB, Canada, 4–8 July 2004; p. 78.
52. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman & Hall: Boca Raton, FL, USA; New York, NY, USA; Washington, DC, USA; London, UK, 1993.
53. Ho, T.K. Random decision forests. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Washington, DC, USA, 14–15 August 1995.

54. Stellmacher, G. *Mapping Kibera: New Strategies for Mapping and Improving the Slum*; Technical Report; University of Washington: Seattle, WA, USA, 2011.
55. Liu, H.; Huang, X.; Wen, D.; Li, J. The use of landscape metrics and transfer learning to explore urban villages in China. *Remote Sens.* **2017**, *9*, 365. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).