



**European Journal of Remote Sensing** 

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tejr20

# Litter on the streets - solid waste detection using **VHR** images

Yrneh Zarit Ulloa-Torrealba, Andreas Schmitt, Michael Wurm & Hannes Taubenböck

To cite this article: Yrneh Zarit Ulloa-Torrealba, Andreas Schmitt, Michael Wurm & Hannes Taubenböck (2023) Litter on the streets - solid waste detection using VHR images, European Journal of Remote Sensing, 56:1, 2176006, DOI: 10.1080/22797254.2023.2176006

To link to this article: https://doi.org/10.1080/22797254.2023.2176006

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



d

View supplementary material

đ	1	ſ	}	

Published online: 20 Feb 2023.



Submit your article to this journal 🕝



View related articles



View Crossmark data 🗹

Taylor & Francis Group

Taylor & Francis

OPEN ACCESS OPEN ACCESS

# Litter on the streets - solid waste detection using VHR images

Yrneh Zarit Ulloa-Torrealba D<sup>a,b,c</sup>, Andreas Schmitt D<sup>a,d</sup>, Michael Wurm D<sup>b</sup> and Hannes Taubenböck D<sup>b</sup>

<sup>a</sup>Department of Geoinformatics, Munich University of Applied Sciences HM, Munich, Germany; <sup>b</sup>German Aerospace Center (DLR), Earth Observation Center (EOC), Oberpfaffenhofen, Germany; <sup>c</sup>Chair of Remote Sensing, University of Wuerzburg, Wuerzburg, Germany; <sup>d</sup>Institute for Applications of Machine Learning and Intelligent Systems (IAMLIS). Munich, GermanyDepartment, University, Munich, Germany

#### ABSTRACT

Failures in urban areas' solid waste management lead to clandestine garbage dumping and pollution. This affects sanitation and public human hygiene, deteriorates quality of life, and contributes to deprivation. This study aimed to test a combination of machine learning, high-resolution earth observation and GIS data to detect diverse categories of residual waste on the streets, such as sacks and construction debris. We conceptualised five different classes of solid waste from image interpretation: "Sure", "Half-sure", "Not-sure", "Dispersed", and "Non-garbage". We tested a combination of k-means-based segmentation and supervised random forest to investigate the capabilities of automatic classification of these waste classes. The model can detect the presence of solid waste on the streets and achieved an accuracy of up from 73.95%–95.76% for the class "Sure". Moreover, a building extraction using an EfficientNet deep-learning-based semantic segmentation allowed masking the rooftops. This improved the accuracy of the classes "Sure" and "Non-garbage". The systematic evaluation of all parameters considered in this model provides a robust and reliable method of solid waste detection for decision-makers. These results highlight areas where insufficient waste management affects the citizens of a given city.

#### **KEY POLICY HIGHLIGHTS**

- The best segmentation using simple linear iterative clustering (SLIC) was achieved with the
  parameter values 8,000 segments and 0.3 compactness. The following supervised classification
  of the segmented images using Random Forest yielded an average overall accuracy of 80.18%.
- The model can detect the presence of solid waste on the streets and achieved an accuracy of up from 73.95%–95.76% for the class "Sure".
- The average reflectance values of the classes "Sure" and "non-Garbage" overlapped. Removing the building rooftops from the orthotiles reduced the overlap of the classes mentioned above. This allowed better identification of the class "Sure".
- Moreover, rooftop removal helped improve the accuracy of the classifier, from 59.51% to 90.18% to 71.53% to 95.76% in study areas with and without rooftops, respectively.

#### **ARTICLE HISTORY**

Received 22 June 2022 Revised 11 November 2022 Accepted 31 January 2023

#### **KEYWORDS**

Solid waste; sanitation; deprivation; remote sensing; machine learning; superpixels

### Introduction

Can we monitor garbage on the streets? Can we use remote sensing together with an automatic or semiautomatic method to identify where there are sanitary problems in a city?

Sanitation, a human right, refers to the access to and use of facilities to dispose of solid waste appropriately, among others (Habitat, 2020). Unfortunately, not all institutions or governments have the capacity and resources to provide the necessary services to the residents, like proper sanitation, fast enough (Habitat, 2020). The "management of solid waste and stormwater drainage", also named "*environmental sanitation*" (HABITAT, 2008), affects the individual and the community as well.

At a local scale, municipalities are typically the ones in charge of the collection, transport, and final disposal of

solid waste (HABITAT, 2008). When this service fails, residents might discard their garbage legally or illegally in open spaces, parks, and rivers or leave it accumulating on the streets. This waste can be transported to other areas because of strong winds or rain, polluting other neighbourhoods, rivers, or groundwater. When left on the streets, it clogs the drainage system, causing flooding (Medina, 2010). Improper management of solid urban waste contaminates groundwater (Vasanthi et al., 2008), attracts pests and animals (e.g. rats) that transmit diseases, and contaminates the air, among others (Yang et al., 2018). Since the decomposition of garbage also occurs in an anaerobic way, it produces methane, which in turn causes spontaneous fires. Moreover, some people might induce fires to burn waste and reduce the sanitary impact and the volume of waste in the dumps (Medina, 2010).

**CONTACT** Yrneh Zarit Ulloa-Torrealba Sulloa-to@hm.edu Dunich University of Applied Sciences HM. Munich, Germany Supplemental data for this article can be accessed online at https://doi.org/10.1080/22797254.2023.2176006.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

These challenges of unmanaged waste predominantly appear in cities of the Global South, for example, in Latin America and the Caribbean, and especially in poor urban areas, such as informal settlements, where public services are often not comprehensive (Martínez Arce et al., 2010). With the expected increase in the global urban population to 60.4% by 2030, the number of slums or areas with deprived urban infrastructure will also increase with sanitary problems (Habitat, 2020; Medina, 2010).

The lack of proper sanitation or poor management of urban solid waste deprives the population of basic hygiene and health, leading to a lower quality of life (HABITAT, 2008). This deprivation of basic needs and opportunities limits the individual's ability to live a fulfilling life, thereby enhancing poverty (Anand & Sen, 1997; Kuffer et al., 2018; Taubenböck et al., 2018). To tackle these issues and thereby improve the well-being of urban inhabitants is part of the United Nations Agenda of Sustainable Development Goals (SDGs) objectives. More specifically, targets 6.3 and 11.6.1 aim for the appropriate disposal of waste to avoid the pollution of water sources (UN-Water, 2017) and promote the sustainable management of solid waste (Habitat, 2016). Therefore, identifying areas with deficiencies in sanitation or solid waste management (SWM) can support urban planning and management.

# Research on remote sensing to study waste management in urban areas

Remotely-sensed data can provide information on the location where the garbage was or should be disposed. The use of sensor products varies with the size and characteristics of the solid waste being studied. For example, Gill et al. (2019) used Landsat TM and ETM+ to detect the waste that spread under landfills with a Ground Sampling Distance (GSD) of 100 m. For the monitoring of waste on land areas approximately  $2 \times 2$  m in size, Yonezawa (2009) combined data from ALOS and Quickbird (0.65-2.5 m GSD). Karimi et al. (2022) used Landsat 8 and night light images from the Suomi NPP to estimate the probability of locating illegal landfills (30-500 m GSD). In general, images from high-to middle-resolution sensors can be used for the identification of dumping sites at a scale of a few metres.

In the case of dumping zones a few centimetres in size, very high-resolution (VHR) imagery is necessary. Data from airborne cameras or unmanned aerial vehicles (UAVs) are available in the order of millimetres or centimetres, depending on the flight altitude, quality of the camera, and atmospheric conditions (Osco et al., 2021). Jakovljevic et al. (2020) used UAV (0.4– 2.3 cm GSD) data to detect plastic bottles in water bodies. Torres and Fraternali (2021) used UAVs (20 cm GSD) to detect and map illegal dumping zones. To achieve greater detail on the nature and extent of solid waste, other sources have been used, such as data from photos or images from surveillance cameras (Alfarrarjeh et al., 2018; Dabholkar et al., 2017), a combination of Google Street View, ImageNet, and self-taken images (Ping et al., 2020), or repositories of data like SpotGarbageGINI in GitHub (Patel et al., 2021). In all of these cases, objects like bottles, cartons, furniture, etc., were visible and easily identified.

For solid waste data analysis, several methods have been tested. Visual estimations of dumping areas can be helpful if it is not possible to access them and were the most successful on sites <400 m<sup>2</sup> in Bangalore, India (Chanakya et al., 2017). Diverse machine learning methods, such as deep learning (DL) (Dabholkar et al., 2017; Jakovljevic et al., 2020; Patel et al., 2021; Ping et al., 2020; Torres & Fraternali, 2021; Youme et al., 2021), and decision trees classifiers (Alfarrarjeh et al., 2018), among others (Shahabi et al., 2014) (for a more detailed review, see (Singh, 2019; Xia et al., 2021). However, some studies also relied on spectral signature differences (Yonezawa, 2009) or visual change detection to estimate the dumping zones' location.

Regardless of the extensive research on solid waste identification using remote sensing methods, when urban deprivation in cities is estimated, the waste aspect is integrated using GIS or survey-based methods (Ajami et al., 2019; Kuffer et al., 2021). For example, Ajami et al. (2019) measured the deprivation of a slum using a set of surveyed and remotely-sensed factors. Waste management was only part of the survey (i.e. GIS data). While there is general agreement with this approach, we believe that the estimation of urban deprived areas could also benefit from a remote sensing - based method as a proxy of sanitary deprivation. After all, the more accurate the data for urban areas, especially the ones that struggle the most, the better we can provide information for policymakers, and stakeholders, among others (Kuffer et al., 2021).

#### Solid waste conceptualisation

Illegal waste disposal has different meanings depending on many factors, including the following:

- (1) The legal system (i.e. how does a local government define litter?).
- (2) The components (i.e. domestic, or construction waste, among others).
- (3) Size (i.e. from a few square centimetres to several hundred square metres).
- (4) The behaviour of the citizens (e.g. dumping zones on the streets or around collection centres).

Different locations have laws to define illegal littering. For example, in a study of illegal dumping in Queensland, Australia, the authors stuck to the local legislation for a definition of the type of waste on which their research focused: "illegal waste disposal sites are restricted to the unlawful deposit of an amount of domestic waste 200 litres or greater in volume" (Glanville & Chang, 2015). In Colombia, Law 120-99 of the National Congress states where solid waste should not be disposed of and, if so, how the person should be penalised (Congreso Nacional, 1999). Although the law does not define solid waste, it states that garbage should not be disposed of on "streets, sidewalks, curbs, parks, highways, roads, public baths, seas, rivers, creeks, streams, and irrigation channels, beaches, squares and other places of recreation and other public places" (Congreso Nacional, 1999).

The size of the dumping zones depends on their characteristics. Generally, solid waste refers to objects or materials that are useless to humans and, therefore, discarded (Medina, 2010). Waste can be divided into several categories: household solid waste, municipal, or urban solid waste, special waste, construction waste, and hazardous waste (Martínez Arce et al., 2010). Waste is defined by sources, such as the households of city residents, generated during production processes, produced by the construction or demolition of infrastructure, or by activities that could affect human health. These can be in solid, liquid, or gaseous form (HABITAT, 2008; Martínez Arce et al., 2010; Xia et al., 2021). Even though there is research on clandestine littering on streets using remote sensing and or artificial intelligence (AI) methods, all studies have diverse definitions of garbage (Alfarrarjeh et al., 2018; Dabholkar et al., 2017; Patel et al., 2021; Ping et al., 2020; Torres & Fraternali, 2021).

When waste is packed in plastic bags, regardless of content, specific elements are escape the remote sensing detection. With VHR or camera surveillance imagery, it is possible to detect specific elements like furniture, electronics (Alfarrarjeh et al., 2018; Dabholkar et al., 2017), or plastic bottles (Jakovljevic et al., 2020). Detection focused on garbage bag accumulation or small piles of litter on the streets might be useful, especially for low-income countries that struggle with their SWM (Iyamu et al., 2020), and when VHR imagery is not available for detecting individual objects.

In this study, we developed a model for detecting solid waste that focuses on objects disposed on the streets or areas of public access that are not dumped into containers but instead abandoned, cornered, or grouped into visually defined clusters. Usually, these waste objects are packed into white or black bags, creating compact objects that can be recognised in several locations. For this purpose, we defined classification categories based on the probability that an object was garbage.

As a case study, we focused on Medellín, Colombia. Local media constantly reports about citizens dumping their waste on the streets outside the containers designated for its disposal. The municipality struggles to identify the more affected zones and the citizens who litter illegally (El Tiempo, 2022). The novelty of this work is to develop a model of solid waste identification focused on aggregations of litter (like bags) dumped in streets or areas of public access and not dumped into containers or landfills, which have been the main focus of most of the recent studies in this topic. Moreover, our model uses imagery provided by the local government of Medellín, which allows for faster implementation of SWM programmes.

#### **Objectives**

This study aimed to test a combination of remote sensing data and machine learning approaches to conceptualise and detect illegal solid waste dumping in an urban landscape. In this way, we can provide a reliable method to decision-makers on where insufficient waste management affects urban residents. Described below are the steps of the workflow:

- (1) Supervised segment-based classification of orthorectified images to detect urban waste accumulations.
- (2) Evaluate which appearance or type of urban waste can be detected at which accuracy levels with the approach mentioned above.
- (3) Determine if an auxiliary data set on the buildings improves the capacity to identify street waste accumulations.

In the following chapters, we (i) describe the utilised materials and explain the developed methods, (ii) focus on the results of our experiments, and (iii) explain the outcome of our analysis and implications for policymakers or decision-makers.

#### Materials and methods

The following section describes the datasets and the algorithms used.

#### Study area and data

The research used data from Medellín, Colombia (Figure 1). This municipality belongs to the Department of Antioquia. Its authority extends over 374.8 km2, which contains 16 communes and 273 neighbourhoods in 117.4 km2. The working area or region of interest (ROI) comprises 23.04 km2. This is



Figure 1. Relative location and overview of the study area, Medellín. The administrative borders of the different neighbourhoods are shown in green.

defined by the area covered by the available aerial images. From this ROI, 25 areas of interest (AOI) were selected, each 0.25 km2. These AOIs are image subsets in which the analyses were conducted.

The research data utilised in this study include raster and vector datasets: (1) orthorectified aerial images, (2) building and rooftop footprint, and (3) labelled polygons (Figure 2).

The orthorectified images comprise four bands (blue, green, red, and near-infrared) with a pixel size of 8 cm. Each image is a composite created by the mosaic of different stripes of camera recording underneath an aeroplane (Servicios de Imágenes de Medellín, 2021) and covers an area of 3.84 km<sup>2</sup>. All images were reprojected to the Antioquia Medellín coordinate system with Datum MAGNA and Mercator Projection. The images were from 2019 and were provided by the Image Service of the Municipality of Medellín via an ArcGIS online server (Servicios de Imágenes de Medellín, 2021).

Covering the entire city of Medellín, building footprint data outlined the borders of all buildings with rooftops. The dataset from 2017 is provided by the GeoMedellín Service of the Municipality of Medellín (GeoMedellin, 2020). Since this dataset did not include all buildings created from 2017 onwards, an updated building footprint data set was created based on semantic segmentation of the orthophotos (see the Building Footprint section) and merged with the official one. Since garbage is usually found on the streets, rooftops are excluded by masking out the building footprints. This allowed the classifier to focus on the streets. The intent was to detect garbage that poses a hygiene risk to the urban population. Therefore, any element that resembled garbage inside a private property was beyond the scope of this research. The training and test data included labelled polygons in two main categories:

- (1) Areas that included garbage or urban residual waste (G).
- (2) Areas with anything else that was not garbage (nG).

The G-dataset was created manually using visual recognition of the garbage accumulation on the



**Figure 2.** Vector and raster datasets used in this project: a) shapefiles clipped to the extent of the 25 AOIs, b) shapefile of building and rooftop footprint, and c) orthotiles, 8 cm GSD, 2019, clipped to the ROI.

orthorectified images of the ROI and their posterior mapping. This process produced a total of 2,660 training areas for detected waste. The creation of the nG dataset was done using segmentation (see the section on Sensitivity Analysis and Segmentation) and posterior selection of segments representing the diversity of nG elements on the 25 AOI raster images. For the input dataset, 500 samples of G and nG objects were randomly selected. Finally, this input dataset was split into 70% train and 30% testing.

## **Building footprint**

The official building footprint data set provided by the city of Medellín dates from 2017, while the orthophotos for garbage detection date from 2019. The entire orthophotos, at 16 cm pixel size, were split into individual image tiles of  $224 \times 224$  pixels with 33% overlap between the images to reduce border effects. Using this high-resolution remote sensing data, we created an updated building footprint dataset with a deeplearning-based semantic segmentation approach (Wurm et al., 2019, 2021) (Figure 3). A precise and complete data set on building footprints was essential for the success of the presented method for garbage detection.

For the process of building extraction using semantic segmentation, we used EfficientNet, as introduced by Ronneberger et al. (2015). One of the main advantages of this architecture is that it can deal with a small number of samples, which is advantageous in the context of building extraction. Furthermore, the network uses data augmentation to artificially increase the number of training samples. The network was trained with local domain knowledge from Medellín Orthophotos with manually derived building footprints. Detailed information on the model set-up and parameters can be found in (Wurm et al., 2021).

# Urban residual waste dataset

Samples for training the model were created by visual interpretation. This resulted in more than 3,000 polygons assigned to five different categories, indicating the reliability of the objects being garbage or not. These are termed "Sure", "Half-sure", "Not-sure", "Dispersed", and "Non-garbage" (Figure 4). The class "Sure" was composed of grouped black-and-white round-shaped objects easily recognisable as garbage elements. Comparing some "Sure" locations with Google Street View data confirmed that it refers to bags piled up and disposed of on the streets.

The classes "Half-sure" and "Not-sure" refer to the 50% and<25% probability that the selected objects are garbage, respectively. These proportions are based on ground-based user experience, i.e. some of the authors have observed the littering problem in the region. Small, scattered elements of garbage covering an empty area or with ground surface visible in between are labelled as "Dispersed". They are probably garbage, but they are not packed into bright and black bags like the "Sure" ones. Everything else in the scenes that do not belong to the categories above was labelled "Non-garbage", i.e. dwellings, streets, humans, vehicles, vegetation, and rivers, among others.

To determine which residual waste classes could be identified, all waste categories were combined with nG polygons. The five categories mentioned above were combined in different ways, called "Treatments" in this paper. Treatments refer to the five combinations of the solid waste factor on which the classifier is applied. They were defined as follows (see also Figure 4):

- A: Sure + nG
- B: Sure + Half-Sure + nG
- C: Sure + Dispersed + nG
- D: Sure + Half-Sure + Not-Sure + nG
- E: Sure + Half-Sure + Not-Sure + Dispersed + nG

In conclusion, based on the ground sampling size and temporal frequency of the orthophotos, theoretically it is possible to detect solid waste objects of at least 256  $cm^2$  when the aeroplane camera is recording. However, the segmentation algorithm will also influence the final minimum size, since it clusters the pixels in searching for homogeneity zones. Our waste objects



Figure 3. A subset of an orthotile: a) with buildings and b) without building footprints.

have a minimum size of  $760 \text{ cm}^2$ . Since it is impossible to detect individual elements such as plastic bottles or cartons with 8 cm GSD, we defined the probability mentioned above as classes of garbage. The categorisation of waste was based on the user ground-based experience. Finally, the VHR imagery available is recorded once per year. This only allows for the detection of waste at a specific time of the year.

### Sensitivity analysis and segmentation

SLIC is an unsupervised k-means-based algorithm for spatial image segmentation. This classifier groups or clusters pixels based on their colour similarity and closeness on the image plane, thereby reducing the complexity of an image (Achanta et al., 2010). It is based on a five-dimensional space named CIELAB, composed of the *Labxy* parameters: [Lab], which represents the colour vector using lightness L and chromaticity ab, and [xy], which represents the location coordinates of a given pixel (Achanta et al., 2010, 2012). Segmentation results in a series of superpixels or regions of homogeneity. These superpixels are a set of pixels grouped into a segment that does not necessarily represent a semantic object completely, but a homogeneous part (Ren & Malik, 2003).

The implementation of the SLIC algorithm in Python was accomplished via the skimage library (van der Walt et al., 2014). The SLIC function assigned each pixel i to the closest cluster, and in every iteration, the distance was reduced (Achanta et al., 2012). The Python implementation of the SLIC function



**Figure 4.** Training data categories were defined in this study. There were four categories of urban residual waste":sure"",halfsure"", and not-sure", which represent 100%, 50%, and<25% probability of being garbage, respectively, as well as "dispersed", and non-garbage. The first row shows the orthotile subsets, while the second row shows the same location in google street view.

clustered the pixels based on the parameters "number of segments" (ns) and "compactness" (c), where the first defined the approximate number of segments to fit in the image. The latter measures the compromise between colour and spatial proximity (van der Walt et al., 2014).

The goal was to segment each of the 25 AOIs so that the segments could capture even the smallest garbage areas. To determine which parameter values would generate meaningful segmentation in our AOIs, a supervised sensitivity analysis was applied. After a brief inspection of the SLIC function and how it performs with our data, we systematically tested the following values for the "number of segments": 2,000, 4,000, 6,000, 8,000, 10000, 12000, and 14,000. The following values were tested for "compactness": 0.001, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1. This analysis was performed on the randomly selected AOIs 1, 25, 6, 14, and 17, resulting in 560 segmentations.

Figure 5 illustrates the effect of different compactness values for a fixed number of segments (8,000 ns). The lower the compactness, the larger and less homogenous the segments. For compactness>1, the quality rate (QR) and the number of squared polygons stabilise until the values for both functions do not change further (for more detailed information, see Supplementary Material, Tables 1 and 2). After



Figure 5. The subset of an AOI segmented with the SLIC algorithm. The parameter values tested were as follows: number of segments = 8,000, compactness = [0.0001, 0.001, 0.01, 0.1, 1].

running all the segmentations, the resultant polygons were tested against the digitised garbage dataset. An accuracy metric and a measure of shape were used to estimate the *goodness* of the segmentation algorithm (Clinton et al., 2010).

QR was chosen as an accuracy metric because it considers false-positive errors. In this way, we can avoid segment sizes that cannot fit into a potential garbage object but are incorrectly identified as suitable for the task (i.e. we want to avoid a Type I error)

**Table 1.** Selected values from the sensitivity analysis using the sum of the QR and the proportion of polygons that are not squares (NSP) in each vector scene. NS is the number of segments, and C is compactness.

-				
NS	С	QR	NSP	Sum
10000	0.1	0.266	0.998	1.264
8000	0.3	0.278	0.986	1.264
12000	0.1	0.29	0.997	1.287

		OA		Кар	ора
NS – C	Т	wB	woB	wB	woB
8000-0.3	А	88.52	95.76	0.62	0.69
	В	90.18	83.51	0.65	0.6
	С	83.56	78.9	0.5	0.49
	D	69.83	73.92	0.35	0.51
	E	65.19	72.41	0.31	0.46
10000-0.1	Α	79.62	92.6	0.4	0.75
	В	82.03	81.75	0.53	0.45
	C	72.83	84.83	0.29	0.5
	D	78.04	71.53	0.46	0.36
	E	59.71	76.54	0.32	0.48
12000-0.1	Α	82.04	88.62	0.45	0.68
	В	73.95	80.57	0.33	0.52
	C	66	79.31	0.27	0.53
	D	69.22	73.35	0.31	0.38
	E	59.51	77.44	0.31	0.35

Table 2. Overall accuracy (OA) and kappa for all treatments (T), number of segments (NS), compactness (C), and presence (wB) or absence (woB) of the building footprint.

(Bramer, 2016). QR is an area-based measure that calculates the proportion between the intersection and union of digitised and algorithm-segmented polygons (Clinton et al., 2010; Weidner, 2008). In equation 1, *R* refers to the polygons of the reference dataset, and *S* to the segments created by the SLIC segmentation that we want to evaluate. QR takes values with  $\rho_q \in [0,1]$ , with 1 being the optimal segmentation (Weidner, 2008).

$$\rho_{q} = \frac{|S \cap R|}{|S \cup R|} = 1 - \frac{|(S \setminus R) \cap (R \setminus S)|}{|S \cup R|}$$
(1)

The SLIC segmentation produced squared segments for high values of the compactness parameter. In this case, homogeneity no longer plays a role (Figure 5). Since we wanted segments that considered the influence of the spectral signature, we excluded the polygons with four right-angle vertexes using formula 2:

$$P = 4\sqrt{A_i} \tag{2}$$

where *A* equals the area of each *i* segment. For every combination of ns and c, if *P* equals the perimeter, it is considered a "perfect square" or a segment with four right-angle vertexes. The proportion of polygons *X* that are not squares was calculated with X = 1 - P. Both indexes, *X*, and  $\rho_q$ , were summed. For the classification step, three of the highest values of *X* were selected (Table 1).

#### Factors and classification

To test if the garbage accumulations could be detected in aerial imagery with machine learning approaches,



**Figure 6.** Workflow: A sensitivity analysis of six AOIs was performed, and the building footprint was created. The input data preprocessing involved the following datasets and steps: a) five orthotiles, b) 25 AOIs, c) three SLIC segmentation parameter combinations resulting from the sensitivity analysis, d) five treatments or combinations of garbage categories, and e) two building conditions with buildings or without buildings footprint. All possible combinations of factors [b:e] were calculated. Six statistical metrics were calculated and integrated into the model for each segment. The resultant input dataset was divided into training and testing. This model was evaluated for a total of 750 classifications; later, the accuracy was estimated.

a combination of many framework conditions or *factors* were chosen, namely (Figure 6):

- 25 AOIs in scene subsets taken from five orthotiles.
- Three appropriate sets of parameter values for SLIC segmentation (see Table 1).
- Five combinations (A E) of garbage categories or treatments.
- Two building conditions: Each raster subset or AOI was classified as a whole or without buildings. In the second case, we used a building footprint to clip the rooftops of dwellings out of the scene.
- Six statistical metrics: For each segment, six metrics were calculated for each of the four bands separately. These are minimum pixel value, maximum pixel value, mean, variance, skewness, and kurtosis. The metrics provided a higher-dimensional feature space for classification than only the pixel values.

The combination of the factors mentioned above produced 750 classifications. Each subset was classified using a random forest (RF) approach. RF is an algorithm that performs a classification using decision trees. This classifier was trained with samples selected randomly and with replacement. Roughly two-thirds of these samples were used to create decision trees, while the remaining third was used to validate these trees, in other words, to measure the model's accuracy. The model must define two variables: the number of decision trees and the number of variables to be used when making every decision that leads to a tree. For this study, RF was chosen for the following reasons: its high computational speed and accuracy, it does not assume a normal distribution, and its implementation is quite simple, since it requires only setting up two parameters (Belgiu & Drăguț, 2016; Breiman, 2001). It has been successfully used in a wide range of remote sensing data, from low to VHR images, in combination with other products, to detect land cover classes (for an overview, see (Belgiu & Drăguț, 2016)).

RF was implemented using the "RandomForestClassifier" function from the *Scikitlearn* package in Python (van der Walt et al., 2014). The model was applied using 500 trees and bootstrapped samples. Using the training data, the model was fit using the metadata of the segments belonging to each of the classes mentioned above and their corresponding labels.

#### Accuracy assessment

The estimation of the classification accuracy was grouped into different categories: the segmentation values used, the type of waste, and the presence or absence of building footprint. A confusion or error matrix per group was calculated, and the following indexes were measured: overall accuracy (OA), producer's (PA), user's accuracy (UA), and kappa coefficient. Furthermore, PA and UA were summarised in the F-score for better readability of the results. The F-score is defined as the "harmonic mean between precision P and recall R" (Dalianis, 2018), or PA and UA, and it is defined by formula 3:

$$F - score: F_1 = F = 2 * \frac{P * R}{P + R}$$
 (3)

The confusion matrix summarises how the samples from the test dataset correspond to the categories of the same pixels from the classified image (Bramer, 2020). The correctly classified pixels, related to the total number of pixels evaluated, correspond to OA (Congalton, 2001). The kappa coefficient measures the agreement between the classified image and the reference data and has a value range of [-1,1]. The closer the value is to 1, the higher the agreement between the classified image and the reference dataset (Congalton, 2001). To get an idea of the performance of each class, PA and UA were calculated (Story & Congalton, 1986). The PA measures the "errors of omission" or the probability of a class being correctly classified - in other words, how well the algorithm predicted every class. On the other hand, the UA measures the "errors of commission" or the probability that the classification is what, in reality, is happening in the area studied - in other words, how reliable it is (Congalton, 2001; Story & Congalton, 1986).

#### Results

In the following section, we describe in detail the results and accuracy metrics for the various steps of the workflow. We found that small and heterogeneous urban residual waste can be identified in VHR aerial images using an RF classifier with high accuracy.

#### **SLIC** segmentation

For the sensitivity analysis of the SLIC segmentation, the higher the *ns* and *c* values, the higher the QR (0.017 to 0.354; see Supplementary material). Segmentations with<8,000 ns produced very big superpixels. For example, polygons with [2,000 ns, 0.001c] and [4,000 ns, 0.001c] were huge and yielded lower QR values. These values are located in the upper left light region of the heatmap in (Figure 7). Since vast segments are not suitable for detecting garbage polygons, they were excluded.

On the other hand, high compactness levels affected the shape of the produced segments. The estimation of the proportion of square polygons per scene showed that approximately 99% of the segments



**Figure 7.** Heatmap of the sensitivity analysis showing the combination of the QR with the proportion of non-square polygons. The upper left corner represents segments with low QR and a high proportion of non-square polygons. The bottom right corner represents segments with high QR and a low proportion of non-square polygons. The optimum values are shaded in dark orange.

with  $c \ge 10$  were mostly perfect rectangles. In this case, the shape and homogeneity information of every superpixel were lost. Moreover, the accuracy of the segmentations>12,000 ns and>10c was the same, which means the function reached a plateau after these values. These QR and shape values were primarily found in the lower right diagonal of the heatmap (Figure 7). After discarding all non-suitable values, three values were selected (see Table 1). These three combinations of SLIC function parameters were used to segment the 25 AOIs for the analysis in this study (Figure 8).

The choice of the values of the SLIC segmentation parameters, ns, and c, influenced the classifications. In

general, the segmentation 8,000 ns - 0.3c performed the best in terms of OA. The average OA of the segmentation 8,000 ns - 0.3c was 80.18%, followed by 10,000 ns - 0.1c with an OA of 77.95%, and 12,000 ns - 0.1c with an OA of 75% (see detailed values in Table 2).

## **Building footprint**

The accuracy of the resulting updated building footprint was evaluated using official cadastral building data, yielding 80% accuracy. Specifically, this building footprint had an accuracy of F1: 0.92, precision: 0.89, and recall: 0.94.



**Figure 8.** Selected SLIC parameter values: (a) 10000 ns - 0.1c, (b) 8,000 ns - 0.3c, (c) 12000 ns - 0.1c. A digitised garbage object is delineated in yellow.

Overall Accuracy with - without building mask



**Figure 9.** Barplot of the overall differential accuracy concerning the presence or absence of a building footprint. The data show one bar per segmentation. Positive values mean that OA is higher after clipping out the rooftops. Negative values indicate that the OA was higher when the classification was performed with buildings included. Bars are grouped by treatments, namely: A":Sure" +"non-garbage", B":Sure" +"half-sure" +"non-garbage", C":Sure" +"dispersed" +"non-garbage", D":Sure" +"half-sure" +"not-Sure" +"non-garbage", E":Sure" +"half-sure" +"non-garbage", "non-garbage", D":Sure" +"half-sure" +"non-garbage", "non-garbage", "non-garbage", "non-garbage", D":Sure" +"half-sure" +"non-garbage", "non-garbage", D":Sure" +"half-sure" +"non-garbage", "non-garbage", "no

Adding a building footprint increased OA on almost all treatment and segmentation combinations (Figure 9). This graph shows the difference in the OA minus with the building footprint. The OA is generally



**Figure 10.** F1 score for all treatments [A–E] and segmentations (white/grey shades), with and without the building footprint. A":Sure" +"non-garbage", B":Sure" +"half-sure" +"non-garbage", C":Sure" +"dispersed" +"non-garbage", D":Sure" +"half-sure" +"not-Sure" +"non-garbage", E":Sure" +"half-sure" +"not-sure" +"not-sure" +"non-garbage".

higher in the results where the classification is limited to areas outside building rooftops (from 71.53% to 95.76%) (Table 2). Otherwise, OA was lower (from 59.51% to 90.18%).

A closer look at every class shows that the F1 score tends towards being higher without the building footprints, especially for "non-Garbage", "Sure", "Dispersed", and "Not-Sure" (Figure 10). However, the class "Half-Sure" often performed better with the presence of rooftops (for treatments B and D).

When the buildings were removed, the algorithm located residual waste objects, mainly on the sidewalks where the garbage is usually dumped. Moreover, the classifier seldom identified residual waste objects in open areas, such as in the middle of streets, rivers, and vegetation. This confirms the plausibility of our classification results.

#### Identification of urban residual waste categories

The algorithm and combination of different factors evaluated in this study can separate the defined solid waste classes from the nG segments in the selected study areas. However, differentiating the diverse classes proved to be a more difficult task. The detection of urban waste was successfully achieved with the probability class "Sure", and treatment A was followed by B. Treatment A had an OA ranging from 79.62% to 95.76%, whereas the latter had an OA ranging from 73.95% to 90.18%. On the other hand, the "non-Garbage" category scored the highest UA and PA in all treatments (>70%) (Figure 10). In other words, the absence of solid waste was the most accurate result obtained.

The models compares several combinations of the waste and non-solid waste classes. (Table 3) shows the

Table 3. Estimated error matrix of a classification with 8000ns and 0.3c for all garbage and non-garbage classes. Overall accuracy (OA), user accuracy (UA), and producer accuracy (PA) calculations are included. The upper part of the table shows a detailed information of every index per class. The lower part of the table shows a summarized information for all solid waste classes.

	pred./ref.	HS	S	D	NS	nG	UA [%]
Half-sure	HS	1291	387	237	618	384	44.26
Sure	S	719	1564	0	470	0	56.81
Dispersed	D	539	585	460	579	1325	13.19
Not-sure	NS	908	1439	0	1308	1690	24.47
Non-garbage	nG	523	493	219	1079	27388	92.21
	PA [%]	32.44	35.00	50.22	32.26	88.96	OA = 72.42
Garbage	HS, S, D, NS	3457	3975	697	2975	3399	76.56
Non-garbage	nG	523	493	219	1079	27388	92.21
	PA [%]	86.86	88.97	76.09	73.38	88.96	OA = 87.08

confusion matrix of the classification of the 25 AOIs using segments created with the SLIC parameters: 8000ns and 0.3c. "Non-garbage" is the class best identified on the orthotiles with UA = 92.21% and PA = 88.96%. In this example, the class "Sure" with UA = 56.81%, is the second most reliable one. The model can identify and differentiate it from "Non-garbage" and "Dispersed". Approximately 50.22% of the objects classified as "Dispersed" were correctly identified, but only 13% of the scattered solid waste piles found were actually "Dispersed". "Not-sure" and "Dispersed" showed the lowest UA values, which indicates high commission errors, or segments wrongly classified. Finally, the classifier performs best, when all solid waste samples vs non-garbage are evaluated.

The classification results are influenced by the choice of methods and the properties of the data. Therefore, we explored the average spectral signature of each class. The "Half-sure" class exhibits, on average, the highest spectral reflectance values in all bands (Blue: 146.49 ± 35.55, Green: 160.36 ± 32.65, Red: 161.48 ± 30.69, NIR: 108.98 ± 32.66). In addition, the reflectance distribution was easily differentiated from the signatures of all other classes. On the contrary, the average spectral signature of class "Sure" did not overlap with the nG classes only when the building footprint was masked. The average values of the classes "Dispersed" and "Not-Sure" were similar, overlapping in the spectral signature. "Non-Garbage" elements have lower reflectance values in the training data without building footprints. In general, the spectral signature that was best differentiable from the nG datasets was from the class "Half-sure", followed by "Sure" without the influence of the rooftops (Table 4). For more information, see (Figure 1) of the supplementary content.

#### Discussion

In our experiments, we proved it was possible to locate residual waste on urban roads in high-resolution aerial imagery. This was possible with an accuracy of up to 80–90% for class "Sure" when the rooftops were masked, although objects in the class "Half-Sure" were also detected with the entire scene.

# Effects of segmentation on classification

The sum of QR and the rate of non-square polygons shown in (Figure 8) combine the spectral and shape information in one index. The rate of nonsquare polygons allowed segmentations with high QR to be excluded, but with compactness so high that homogeneity no longer played a role. Since solid waste objects do not always have the same appearance and size, the selected values were optimal for the segmentation process. The selected segmentations produced small, not square, segments, and fit into the training garbage areas. Many small segments combined had more chances to overlap with any possible garbage object, increasing the possibility of identifying any shape and size of the garbage area in an image.

The selected parameter values ns and c provide the best balance among all the variables defining the shape and size of superpixels that can detect garbage objects or parts of them. However, it is essential to highlight

Table 4. Overall mean and standard deviation of the spectral signature (reflectance) of each classification category for all four bands. wB: nG dataset with buildings; woB: nG dataset without buildings.

	Blue		Green		Red		Near Infrarred	
Training category	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Sure	84.59	26.22	100.67	24.65	105.95	22.76	56.76	21.52
Half-sure	146.49	35.55	160.36	32.65	161.48	30.69	108.98	32.66
Not-sure	113.2	44.31	126.11	41.4	129.85	38.65	79.75	37.44
Dispersed	116.55	30.01	124.7	26.19	118.86	24.71	91.68	25.34
Non-Garbage – with Building	82.71	39.43	97.62	34.29	99.78	33.62	60.02	29.37
Non-Garbage – without Building	71.01	36.91	89.16	31.88	91.33	31.28	55.56	29.19

that these chosen values are specific to our data for several reasons: the size of each AOI raster (i.e. 0.25 km<sup>2</sup>), the size, shape, and spectral properties of the objects to be found on the image, the spectral information and resolution (i.e. bands red, green, blue, and near-infrared), and the pixel and ground sampling size (i.e. 8 cm pixel size), as well other factors that affected the scenes, such as differences in light intensity or quality, errors in the mosaicking, or shadows.

Nevertheless, by applying a sensitivity analysis using the SLIC algorithm, we developed a systematic and non-subjective method to overcome these challenges and choose the correct values for the final segmentation. Here, we highlight the importance of evaluating other parameter values like the second and third best instead of only the first. Since urban waste objects present high spectral variability, this approach increases the chances of detection.

#### Effects of the building footprint on classification

Training the model without the building footprint allowed for improving the classification for many reasons. First, after removing the buildings, we identified waste in the areas of our public spaces. Second, the remaining area has fewer land cover classes available. A visual inspection of the images without the building footprint indicated that streets, vegetation, bare soil, and water were primarily present. Finally, the spectral information of the study areas with and without rooftops was very different (Table 4). The class "Sure" signature overlapped with "non-Garbage" when rooftops were included. This could be due to the colours of the rooftops, which resemble garbage elements. Therefore, removing the building footprint allowed better identification of the class "Sure". The other solid waste class features have an average spectral signature higher than nG, making them easier to differentiate.

The fact that the accuracy of classes "Half-Sure", "Not-Sure", and "Dispersed" was not always improved when masking the rooftops could have different explanations:

- (1) How the class was defined, or how the objects were assigned to this class.
- (2) How the training data for these classes were created because these objects were not easily identified as solid waste as the class "Sure".
- (3) The spectral signatures of "Not-Sure" and "Dispersed" presented a high overlap. Another feature space could be considered in future analysis; for example, the inter-channel correlation could enhance the distinction between classes.

(4) The algorithm might perform better in identifying single classes than mixtures of them. A future step would be to evaluate those classes independently (similar to Treatment A).

### Identification of garbage and non-garbage areas

The model was successful at identifying what is not garbage, as well as the category "Sure". Visually speaking, the class "Sure" was very homogenous because it was primarily composed of the same types of objects or plastic bags. Hence, objects containing diverse elements, not packed in the usual white – black plastic garbage bags, scored lower accuracy. The UA was mostly higher than the PA, indicating how many segments were identified as "Sure" waste that genuinely belonged to this category of garbage (Congalton, 2001). Other classes and treatments scored lower in UA and PA, which denoted how difficult it was to distinguish them from nG or other classes.

When classes such as "Dispersed" and "Notsure" were included in the treatments, the accuracy dropped. These classes were the most difficult to detect and classify correctly. This could be due to the nature of the objects, i.e. the semantic information used to label those elements on the streets as one class or the other. Removal of these classes can still identify the typical litter, clearly wrapped in bags, dumped along the streets or piled against an electricity pole.

The average reflectance of a band of these categories overlapped significantly in blue and green (Table 4). Including band combinations of blue and green could be a way to identify these classes. The fact that the algorithm performed worse when including these classes could be due to the identification of the elements that belonged to the class itself, to the algorithm chosen, or to the variability of the spectral signature of that class. Difficulties distinguishing solid waste from bare soil have been previously reported. Yonezawa (2009) struggled to identify garbage over the ground without vegetation using multispectral Quickbird data.

During the creation of training data, garbage areas were sometimes challenging to distinguish from other objects on the scene, which were difficult to identify or assign to any class. Sometimes, it was clear that a specific object belonged to a residual waste category, but its appearance differed from other objects of the same class. At other times, objects on the scene looked similar to urban residual waste, for example, motorbikes, car windows, shadows, street drains, heads of pedestrians, or other types of garbage not previously identified. There might also be other solid waste categories found in the images that



**Figure 11.** Detailed view of some examples of the classification performance with the corresponding F1 score. The segments related to the class in the label are highlighted in yellow. Since some classifications involved more than one class, the segments of the other solid waste categories are shown in black. The examples are from the following combinations of treatments, building footprint, and segmentations: a) class"Sure", treatment A, with building footprint, 8,000 ns - 0.3c, b) class"Sure", treatment A, with building footprint, 8,000 ns - 0.3c, d) class"Half-sure", treatment B, with building footprint, 8,000 ns - 0.3c, d) class"Half-sure", Treatment B, without building footprint, 8,000 ns - 0.3c, f) class"Dispersed", treatment C, with building footprint, 8,000 ns - 0.3c, f) class"Dispersed", treatment C, without building footprint, 12,000 ns - 0.1c, g) class"Not-sure", Treatment E, with building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, without building footprint, 10,000 ns - 0.1c, h) class"Not-sure", treatment E, wi

we could not detect due to size or appearance. These features influence the assertiveness of the final classification, which can be seen in the images from (Figure 11). Most misclassifications happened on some of the objects mentioned above, identified as solid waste, and primarily located alongside the streets.

While there were designed centres for waste collection and recycling in the city, our selected AOIs did not overlap. Nevertheless, our model could also identify the garbage inside these locations. If they overlap, these locations could be excluded to generate an accurate view of the illegal dumps. Besides the official centres, there are also authorised locations for solid waste accumulation, which our model can detect. This method can be used to validate illegal dumping zones if combined with ground truth data.

#### Challenges

Due to the spatial resolution of the orthotiles it is only possible to detect objects larger than the sampling size  $(64 \text{ cm}^2)$ . This implies that the model cannot detect small solid waste elements thrown on the streets, such as plastic bottles or cigarette butts. However, when local media reports illegal dumping on the streets,

this includes big plastic bags of domestic waste. Therefore, this model contributes to the detection of a significant component of littering in Medellín.

The temporal resolution of the images of one record per year provides a screenshot of the city. When we apply the waste detection model to these images, we briefly see the city's condition, which might not represent a whole year. Therefore, it is impossible to quantify how much solid waste can be found on the streets of Medellín.

Another aspect of the temporal scale is comparing photos of other dates. The orthotiles are images recorded in 2019, while the Google Street View photos span from 2016 to 2021. Comparing the identified locations with Google Street View did not necessarily indicate that the dumping zones were permanent. However, if certain spots were visible on different time stamps, this might indicate that some illegal dumping occurred regularly, as the local media reported. Nevertheless, the model can detect solid waste, and if more images are available, more accurate temporal quantification could be possible.

The definition of the classes might have affected the capacity of the classifier to identify them. As previously mentioned, sometimes the waste objects looked similar to other elements in the scene. We focused on garbage dumped in large plastic clusters because the average citizen discards all types of garbage in one bag, regardless of which category it belongs (Peralta Miranda et al., 2019). Therefore, the classes were designed on the probability of being garbage and not on its content.

Further research on this topic can benefit from many lessons learned, such as, focus on a garbage vs non-garbage classification, include multi-temporal information, or testing other AI algorithms. More training and adequate data are necessary for better classification performance. For example, Thung and Yang (2017) first built a dataset of images called TrashNet before training a model. The final algorithm helped identify and sort trash in a recycling process. More extensive and diverse training data on solid waste could also improve the detection of dumping zones.

Recently, more studies have been conducted in the SWM field using DL. As an artificial neural network method, it can handle unbalanced or incomplete datasets (Abdallah et al., 2020). This would be suitable for our case study, since most elements on a scene are not solid waste. This could contribute to excluding objects that can be confounded with garbage because of a similar appearance. DL methods can handle higher amounts of nonlinear, complex data faster. However, they are prone to overfitting and will not necessarily improve accuracy compared to decision tree models like the RF tested here (Abdallah et al., 2020).

### Socioeconomic applications

The model proposed in this study contributes to identifying the areas in which SWM collection may be failing. The model demonstrates that it is possible to detect solid waste wrapped in bags and dumped in urban areas. Furthermore, comparing imagery at different times can show which areas are most affected by littering and how it changes over time. Machine learning approaches are not restricted to this phase. Several authors have also contributed to other phases, such as waste bin detection, collection routing optimisation, waste classification for recycling, model parameters of the composting process, and landfill location, among others (Xia et al., 2021).

The local municipality of Medellín reports recurrent disposal of solid waste at unauthorised sites next to a designated dumping site, or even right after the garbage was collected (El Tiempo, 2022). In other words, the problem is not only the garbage disposed of at unauthorised locations, but also that citizens dispose of it right after the regular municipal collection. Due to the impact of residual waste on health and quality of life (Medina, 2010), people need to dispose of their garbage, regardless of whether the local government has an efficient or effective residual waste management system. Using imagery from different times of the day, this model could also identify the zones where people dump their waste outside authorised times.

Classification of images for solid waste can be applied on aerial imagery, or camera surveillance (Ping et al., 2020). The local municipality of Medellín has recently implemented a machine called "Robocop", which does visual recognition and image classification in real time of camera recordings of people who dumps their waste in unauthorized sites. The machine "speaks" with the citizen, and reports the information to the corresponding office. The results are that people feel discouraged to repeat their behaviour. A further step would be to couple this technology with identified critical dumping zones from remotely sensed data.

Garbage collection efficiency might be related to the socioeconomic level of the district (Galvis Gonzalez, 2016). A visual inspection of the orthotiles used in this model shows that districts categorised as middle class or higher seldom have litter on the streets. The failure to provide an adequate residual waste management service could have many reasons: budget, political will, illegal actions of residents, infrastructure, and terrain, among others. Many slums are located in difficult-to-access areas: streets may be narrow or unpaved, or the location may be very steep, hilly, or far away from the disposal centre (Sliuzas & Kuffer, 2008).

Moreover, the residents usually pay the waste management costs via taxes, or in some cities like Bogotá, through the electricity bill (Medina, 2010). Since many slum dwellers do not pay taxes (Medina, 2010) or are not registered users in the electricity grid service, they do not contribute to this service, which aggravates their quality of life. Poor people are just poorer. The application of this model to the entire city of Medellín can also highlight socioeconomic and political problems. Whether the community is rich or poor, or if dumping is legal or illegal, the disposal of litter in public areas imposes a hygiene threat to all citizens (Du et al., 2021).

The implementation of this method requires the generation and processing of orthophotos. However, once automated, this process can be more economically efficient than humans patrolling the streets. With this approach, we hope to generate more knowledge about ineffective waste management and its solutions.

#### Conclusions

This study aimed to test a combination of methods with a conceptual definition of solid waste to detect residual urban waste in Medellín, Colombia. For this purpose, several possible combinations of residual waste, segmentation, and the presence or absence of a building footprint were tested on orthorectified aerial images of Medellín. The methodology for this study focused on statistical robustness, hence the systematic selection of segmentation parameters, the balanced number of samples, and the evaluation of many combinations of factors.

The research methods applied in this study can identify presence of solid waste. While it struggles to differentiate among the categories of garbage, especially "Dispersed" and "Not-sure", it can detect with high accuracy where objects of "Sure" solid waste are disposed of on the streets. In general, the method proved capable of detecting the random waste littered on the streets, serving as valuable information for decision-makers working to enhance SWM.

Accumulations of residual waste in the urban environment are a known public hygiene problem highly correlated with urban poverty (Medina, 2010). Although this research did not explicitly detect the location of poverty, it contributed a method to determine areas in a city affected by a public health issue. Future research should further develop and confirm these initial findings to provide a proxy for urban poverty based on sanitation using remote sensing data.

#### Acknowledgments

Special thanks to Severin Herzsprung and Christoph Otto for their support in the creation of training data. Furthermore, the authors thank Raphael Tubbesing for providing the building footprint. Likewise, we are grateful to Marta Sapena for the support with the data access. Many thanks to Asja Bernd, César Aponte, Nils Matzner, and Sebastian Briechle for their comments and feedback on the content of this paper. Finally, the authors would like to thank the anonymous reviewers for their voluntary and constructive comments, which helped to improve this article.

#### **Disclosure statement**

No potential conflict of interest was reported by the author(s).

#### Funding

The work was supported by the Munich University of Applied Sciences HM and the German Research Foundation (DFG) through the "Open Access Publishing" program German Federal Ministry of Education and Research,FONA Client II initiative,Inform@Risk Strengthening the Resilience of Informal Settlements against Slope Movements [03G0883A-F].

#### ORCID

Yrneh Zarit Ulloa-Torrealba (b) http://orcid.org/0000-0003-3737-7912 Andreas Schmitt ( http://orcid.org/0000-0002-0232-2666 Michael Wurm ( http://orcid.org/0000-0001-5967-1894 Hannes Taubenböck ( http://orcid.org/0000-0003-4360-9126

#### Data availability statement

The data supporting this study's findings are available from the corresponding author, [YZUT], upon reasonable request (Ulloa Torrealba, 2022).

#### References

- Abdallah, M., Abu Talib, M., Feroz, S., Nasir, Q., Abdalla, H., & Mahfood, B. (2020). Artificial intelligence applications in solid waste management: A systematic research review. *Waste Management*, 109, 231–246. https://doi.org/10.1016/j.wasman.2020.04.057
- Achanta, R., Shaji, A., Smith, K., & Lucchi, A. [Aurélien], Fua, P., & Süsstrunk, S. (2010). SLIC superpixels. EPFL Technical Report 149300. https://infoscience.epfl.ch/ record/149300
- Achanta, Radhakrishna, Shaji, Appu, Smith, Kevin, Lucchi, Aurelien, Fua, Pascal, & Süsstrunk, Sabine. (2012). Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11), 2274–2282. https://doi.org/10.1109/TPAMI.2012.120
- Ajami, A., Kuffer, M., Persello, C., & Pfeffer, K. (2019). Identifying a slums' degree of deprivation from VHR images using convolutional neural networks. *Remote Sensing*, 11(11), 1282. https://doi.org/10.3390/rs1111282
- Alfarrarjeh, A., Kim, S. H., Agrawal, S., Ashok, M., Kim, S. Y., & Shahabi, C. (2018). Image classification to determine the level of street cleanliness: A case study. In 2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM) (pp. 1–5). IEEE. https:// doi.org/10.1109/BigMM.2018.8499092
- Anand, S., & Sen, A. (1997). Concepts of human development and Poverty: A multidimensional perspective. In *Poverty and Human Development: Human Development Papers 1997* (pp. 1–20). United Nations Development Programme.
- Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *Isprs Journal of Photogrammetry and Remote Sensing*, 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011
- Bramer, M. (2016). Principles of data mining. Springer London. https://doi.org/10.1007/978-1-4471-7307-6
- Bramer, M. (2020). Estimating the predictive accuracy of a classifier. In M. Bramer (Ed.), Springer eBook collection. Principles of data mining (4th ed, pp. 79–92). Springer London; Imprint Springer. https://doi.org/10.1007/978-1-4471-7493-6\_7
- Breiman, L. (2001). Random forests. *Machine Learning*, 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324
- Chanakya, H. N., Shwetmala, & Ramachandra, T. V. (2017). Nature and extent of unauthorized waste dump sites in and around Bangalore city. *Journal of Material Cycles & Waste Management*, 19(1), 342–350. https://doi.org/10. 1007/S10163-015-0423-6
- Clinton, N., Holt, A., Scarborough, J., Yan, L., & Gong, P. (2010). Accuracy assessment measures for object-based image segmentation goodness. *Photogrammetric*

*Engineering & Remote Sensing*, 76(3), 289–299. https://doi.org/10.14358/PERS.76.3.289

- COHRE, WaterAid, SDC, UN-HABITAT. (2008). Sanitation: A human rights imperative. Centre on Housing Rights and Evictions. https://www.wateraid. org/documents/plugin\_documents/sanitation\_long copy\_eng.pdf
- Congalton, R. G. (2001). Accuracy assessment and validation of remotely sensed and other spatial information. *International Journal of Wildland Fire*, 10(4), 321. https:// doi.org/10.1071/WF01031
- Congreso Nacional. (1999). Ley No. 120-99 que prohíbe a toda persona física o moral tirar desperdicios sólidos y de cualesquiera naturaleza en calles, aceras, parques, carreteras, contenes, caminos, balnearios, mares, ríos, etc. http://extwprlegs1.fao.org/docs/pdf/dom135425.pdf
- Dabholkar, A., Muthiyan, B., Srinivasan, S., Ravi, S., Jeon, H., & Gao, J. (2017). Smart illegal dumping detection. In 2017 IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService) (pp. 255–260). IEEE. https://doi.org/10. 1109/BigDataService.2017.51
- Dalianis, H. (2018). Clinical text mining. Springer International Publishing. https://doi.org/10.1007/978-3-319-78503-5
- Du, L., Xu, H., & Zuo, J. (2021). Status quo of illegal dumping research: Way forward. *Journal of Environmental Management*, 290, 112601. https://doi.org/10.1016/j.jenv man.2021.112601
- El Tiempo. (2022). "Robocop" ya identifica a quienes arrojan basuras en Medellín. https://www.eltiempo.com/colom bia/medellin/medellin-utiliza-robots-para-monitorearbasuras-en-las-calles-698457
- Galvis Gonzalez, José Ariel. (2016). Residuos sólidos: problema, conceptos básicos y algunas estrategias de solución. *Revista GESTIÓN & REGIÓN, 22*, 7–28.
- GeoMedellin. (2020). Portal Geográfico del Municipio de Medellín. Alcaldía de Medellín. https://www.medellin. gov.co/geomedellin/
- Gill, J., Faisal, K., Shaker, A., & Yan, W. Y. (2019). Detection of waste dumping locations in landfill using multi-temporal landsat thermal images. Waste Management & Research: The Journal of the International Solid Wastes and Public Cleansing Association, ISWA, 37(4), 386–393. https://doi.org/10. 1177/0734242X18821808
- Glanville, K., & Chang, H. -. (2015). Remote sensing analysis techniques and sensor requirements to support the mapping of illegal domestic waste disposal sites in Queensland, Australia. *Remote Sensing*, 7(10), 13053-13069. https://doi.org/10.3390/rs71013053
- Iyamu, H. O., Anda, M., & Ho, G. (2020). A review of municipal solid waste management in the BRIC and high-income countries: A thematic framework for lowincome countries. *Habitat international*, 95, 102097. https://doi.org/10.1016/j.habitatint.2019.102097
- Jakovljevic, G., Govedarica, M., & Alvarez-Taboada, F. (2020). A deep learning model for automatic plastic mapping using Unmanned Aerial Vehicle (UAV) data. *Remote Sensing*, *12*(9), 1515. https://doi.org/10.3390/ rs12091515
- Karimi, N., Ng, K. T. W., & Richter, A. (2022). Development and application of an analytical framework for mapping probable illegal dumping sites using nighttime light imagery and various remote sensing indices. Waste Management (New York, NY), 143, 195–205. https://doi. org/10.1016/j.wasman.2022.02.031

- Kuffer, M., Pfeffer, K., & Persello, C. (2021). Special issue "Remote-Sensing-Based Urban Planning Indicators". *Remote Sensing*, 13(7), 1264. https://doi.org/10.3390/ rs13071264
- Kuffer, M., Pfeffer, K., Sliuzas, R., Taubenbock, H., Baud, I., & van Maarseveen, M. (2018). Capturing the urban divide in nighttime light images from the international space station. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(8), 2578–2586. https://doi.org/10.1109/JSTARS.2018.2828340
- Martínez Arce, E., Daza, D., Tello Espinoza, P., Soulier Faure, M., & Terraza, H. (2010). Informe de la evaluación regional del manejo de residuos sólidos urbanos en América Latina y el Caribe 2010. Banco Interamericano de Desarrollo, Organización Panamericana de la Salud, Asociación de ingeniería Sanitaria y Ambiental.
- Medina, M. (2010). Solid wastes, poverty and the environment in developing country cities: Challenges and opportunities. WIDER Working Paper, No. 2010/23 (ISBN 978-92-9230-258-0,). The United Nations University World Institute for Development.
- Osco, L. P., Marcato Junior, J., Marques Ramos, A. P., de Castro Jorge, L. A., Fatholahi, S. N., de Andrade Silva, J., Matsubara, E. T., Pistori, H., Gonçalves, W. N., & Li, J. (2021). A review on deep learning in UAV remote sensing. *International Journal of Applied Earth Observation* and Geoinformation, 102, 102456. https://doi.org/10. 1016/j.jag.2021.102456
- Patel, D., Patel, F., Patel, S., Patel, N., Shah, D., & Patel, V. (2021). Garbage detection using advanced object detection techniques. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 526-531). IEEE. https://doi.org/10.1109/ICAIS50930. 2021.9395916
- Peralta Miranda, P. A., Espinosa Pérez, H., Rico Fontalvo, R., Cervantes, T., Atia, V., & Mulford Hoyos, M. (Ed.). (2019). Comportamiento del consumidor en el manejo de residuos eléctricos y electrónicos en la costa Caribe colmbiana. Barranquilla: Ediciones Universidad Simón Bolívar.
- Ping, P., Xu, G., Kumala, E., & Gao, J. (2020). Smart street litter detection and classification based on faster R-CNN and edge computing. *International Journal of Software Engineering and Knowledge Engineering*, 30(04), 537–553. https://doi.org/10.1142/S0218194020400045
- Ren, X., & Malik, J. (2003). Learning a classification model for segmentation. In *Proceedings/Ninth IEEE International Conference on Computer Vision: Nice, France, October 13 to 16, 2003* (Vol. 1, 10–17). IEEE Computer Society. https://doi.org/10.1109/ICCV.2003. 1238308
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W.M. Wells, & A.F. Frangi (Eds.), Lecture Notes in Computer Science: Vol. 9351. Medical image computing and computerassisted intervention - MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015 : proceedings, part III (Vol. 9351, pp. 234-241). Springer. https://doi.org/10.1007/978-3-319-24574-4\_28
- Servicios de Imágenes de Medellín. (2021). Alcaldía de Medellín. https://www.medellin.gov.co/mapas/rest/ser vices/ServiciosImagen
- Shahabi, H., Keihanfard, S., Ahmad, B. B., & Amiri, M. J. T. (2014). Evaluating boolean, AHP and WLC methods for the selection of waste landfill sites using GIS and satellite

images. *Environmental Earth Sciences*, 71(9), 4221–4233. https://doi.org/10.1007/s12665-013-2816-y

- Singh, A. (2019). Remote sensing and GIS applications for municipal waste management. *Journal of environmental management*, 243, 22–29. https://doi.org/10.1016/j.jenv man.2019.05.017
- Sliuzas, Richard, & Kuffer, Monika. (2008). Analysing the spatial heterogeneity of poverty using remote sensing: Typology of poverty areas using selected RS based indicators. In C. Jürgens (Ed.), Remote sensing - new challenges of high resolution: Earsel joint workshop Bochum (Germany), March 5 - 7, 2008 (pp. 158–167). Selbstverl. des Geographischen Inst.
- Story, M., & Congalton, R. (1986). Accuracy Assessment: A User's Perspective. *Photogrammetric Engineering and Remote Sensing*, 52(3), 397–399.
- Taubenböck, H., Staab, J., Zhu, X., Geiß, C., Dech, S., & Wurm, M. (2018). Are the poor digitally left behind? Indications of urban divides based on remote sensing and twitter data. *ISPRS International Journal of Geo-Information*, 7(8), 304. https://doi.org/10.3390/ijgi7080304
- Thung, G., & Yang, M. (2017). *Classification of Trash for Recyclability Status*. https://github.com/garythung/ trashnet
- Torres, R. N., & Fraternali, P. (2021). Learning to identify illegal landfills through scene classification in aerial images. *Remote Sensing*, 13(22), 4520. https://doi.org/10. 3390/rs13224520
- Ulloa Torrealba, Y. Z. (2022). Data for solid waste detection in Medellín. https://doi.org/10.5281/zenodo.6637715
- UN-Habitat. (2016). Sustainable development goal 11 Make cities and human settlements inclusive, safe, resilient and sustainable: A guide to assist national and local governments to monitor and report on SDG goal 11+ indicators. https://www.local2030.org/library/view/60
- UN-Habitat. (2020). *The value of sustainable urbanization*world cities report: Vol. 2020. UN-Habitat
- UN-Water. 2017. Integrated Monitoring Guide for Sustainable Development Goal 6 on Water and Sanitation: Targets and global indicators. https://www. unwater.org/publications/integrated-monitoring-guidesustainable-development-goal-6-water-and-sanitation. Accessed 23 September.

- van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). Scikit-image: Image processing in python. *PeerJ*, 2, e453. https://doi.org/10.7717/peerj.453
- Vasanthi, P., Kaliappan, S., & Srinivasaraghavan, R. (2008). Impact of poor solid waste management on ground water. Environmental Monitoring and Assessment, 143(1–3), 227–238. https://doi.org/10.1007/s10661-007-9971-0
- Weidner, U. (2008). Contribution to the assessment of segmentation quality for remote sensing applications. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVII(Part B7), 479–484.
- Wurm, M., Droin, A., Stark, T., Geiß, C., Sulzer, W., & Taubenböck, H. (2021). Deep learning-based generation of building stock data from remote sensing for urban heat demand modeling. *ISPRS International Journal of Geo-Information*, 10(1), 23. https://doi.org/10.3390/ ijgi10010023
- Wurm, M., Stark, T., Zhu, X. X., Weigand, M., & Taubenböck, H. (2019). Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. *Isprs Journal of Photogrammetry and Remote Sensing*, 150, 59–69. https://doi.org/10.1016/j.isprsjprs.2019.02.006
- Xia, W., Jiang, Y., Chen, X., & Zhao, R. (2021). Application of machine learning algorithms in municipal solid waste management: A mini review. Waste Management & Research: The Journal of the International Solid Wastes and Public Cleansing Association, ISWA, 734242X211033716. https:// doi.org/10.1177/0734242X211033716
- Yang, H., Ma, M., Thompson, J. R., & Flower, R. J. (2018). Waste management, informal recycling, environmental pollution and public health. *Journal of Epidemiology and Community Health*, 72(3), 237–243. https://doi.org/10. 1136/jech-2016-208597
- Yonezawa, Chinatsu. (2009). Possibility of monitoring of waste disposal site using satellite imagery. *Journal of Integrated Field Science*, 6, 23–28.
- Youme, O., Bayet, T., Dembele, J. M., & Cambier, C. (2021). Deep learning and remote sensing: Detection of dumping waste using UAV. *Procedia computer science*, 185, 361–369. https://doi.org/10.1016/j.procs.2021.05.037