Computers, Environment and Urban Systems 95 (2022) 101820

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Identifying degrees of deprivation from space using deep learning and morphological spatial analysis of deprived urban areas

Angela Abascal [a,b,*], Ignacio Rodríguez-Carreño [c,d], Sabine Vanhuysse [e], Stefanos Georganos [f], Richard Sliuzas [g], Eleonore Wolff [e], Monika Kuffer [g]

[a] School of Architecture, University of Navarra, Pamplona, Spain
[b] Navarra Center for International Development, University of Navarra, Pamplona, Spain
[c] Faculty of Economics, University of Navarra, Pamplona, Spain
[d] Data Science and Artificial Intelligence Institute, University of Navarra, Pamplona, Spain
[e] Department of Geosciences, Environment and Society, Université libre de Bruxelles (ULB), Brussels, Belgium
[f] Division of Geoinformatics, KTH Royal Institute of Technology, Stockholm, Sweden
[g] Faculty of Geo-Information Science & Earth Observation, University of Twente, Enschede, Netherlands

ARTICLE INFO

Keywords:
Remote sensing
Urban footprint
Morphological analysis
GIS
Slums

ABSTRACT

Many cities in low- and medium-income countries (LMICs) are facing rapid unplanned growth of built-up areas, while detailed information on these deprived urban areas (DUAs) is lacking. There exist visible differences in housing conditions and urban spaces, and these differences are linked to urban deprivation. However, the appropriate geospatial information for unravelling urban deprivation is typically not available for DUAs in LMICs, constituting an urgent knowledge gap. The objective of this study is to apply deep learning techniques and morphological analysis to identify degrees of deprivation in DUAs. To this end, we first generate a reference dataset of building footprints using a participatory community-based crowd-sourcing approach. Secondly, we adapt a deep learning model based on the U-Net architecture for the semantic segmentation of satellite imagery (WorldView 3) to generate building footprints. Lastly, we compute multi-level morphological features from building footprints for identifying the deprivation variation within DUAs. Our results show that deep learning techniques perform satisfactorily for predicting building footprints in DUAs, yielding an accuracy of F1 score = 0.84 and Jaccard Index = 0.73. The resulting building footprints (predicted buildings) are useful for the computation of morphology metrics at the grid cell level, as, in high-density areas, buildings cannot be detected individually but in clumps. Morphological features capture physical differences of deprivation within DUAs. Four indicators are used to define the morphology in DUAs, i.e., two related to building form (building size and inner irregularity) and two covering the form of open spaces (proximity and directionality). The degree of deprivation can be evaluated from the analysis of morphological features extracted from the predicted buildings, resulting in three categories: high, medium, and low deprivation. The outcome of this study contributes to the advancement of methods for producing up-to-date and disaggregated morphological spatial data on urban DUAs (often referred to as 'slums') which are essential for understanding the physical dimensions of deprivation, and hence planning targeted interventions accordingly.

1. Introduction

African cities are in a period of rapid urbanization. African urban population grows on average by 3.55% per year (UN-Habitat, 2020). Between 1999 and 2009, the population of Nairobi, Kenya, increased from 2 million to 3.1 million people. The latest census (KNBS, 2019) indicates that the Nairobi urban population reaches up to 4.3 million people, implying an annual growth rate of 3.8% per year. These new urban residents need housing, and if low-cost formal housing is under-provided, many end up living in unplanned areas that grow organically (Davis, 2006; Oberay, 1993). Already, 56% of the urban population in Kenya is living in poor and unplanned areas (World Bank, n.d) referred to as slums by international institutions such as the World Bank (WB) and the United Nations (UN). Specifically, in the city of Nairobi, where

* Corresponding author at: School of Architecture, University of Navarra, Pamplona, Spain.
E-mail address: maabascal@unav.es (A. Abascal).

https://doi.org/10.1016/j.compenvurbsys.2022.101820
Received 2 October 2021; Received in revised form 24 February 2022; Accepted 23 April 2022
Available online 7 May 2022

this study is carried out, 75% of the urban population growth is absorbed by unplanned areas (UN-Habitat, 2006).

There are many ways of referring to unplanned areas with high poverty rates depending on the local context (UN-Habitat, 2003) and these differ between various international organizations, experts (e.g., Sliuzas, Mboup, & de Sherbinin, 2008; UN-Habitat, 2003; UN-Habitat, Statistics-Division, U, and Alliance, C, 2002; World Bank, 2008), and may even vary within the same government itself. For instance, the KENSUP and KISIP programs in Kenya, refer to them as slums or informal settlements. Many historical definitions of slums exist (De Castro, 2018), and although the word slum is commonly used, it has pejorative and ideological connotations (Gilbert, 2007). In contrast, the 'deprived area' designation is considered devoid of ideological meaning given that it is a technical term referring to scarcity. The terminology *deprived urban area* (DUA) is a concept that refers to areas lacking one or more urban indicators at different levels- household level, area level-. Hereafter, the term deprived area will be used to define poor areas and their inhabitants. Although there are efforts to define the indicators that best describe urban deprivation (Abascal et al., 2021; Taubenböck, Kraff, & Wurm, 2018), there is still no global methodological agreement to conceptualise, quantify or map deprivation (Baud, Sridharan, & Pfeffer, 2008; Kohli, Sliuzas, Kerle, & Stein, 2012; Thomson et al., 2020). Identifying and characterizing poverty from a multidimensional perspective is essential for urban planners and policy makers to target and monitor pro-poor policies (Arimah, 2001; Henninger & Snel, 2002; Huchzermer, 2011).

According to UN-Habitat (2003), a "slum" is a household-level concept to characterize deprivation, which is defined as the absence of any one of the following indicators: quality water, improved sanitation, sufficient living space, durable housing, or secure tenure. This definition of a slum is a household-level characterisation that does not incorporate the surrounding area (Lilford et al., 2019). More recently, efforts have been made to understand deprivation as a sum of scarcities and it has been argued that further understanding requires characterisation through a multidimensional perspective (Abascal et al., 2021; Liang, De Jong, Schraven, & Wang, 2021). In line with Thomson et al. (2020), this study argues that encompassing deprivation levels within a broader scope requires not just characterisation of the inhabitants but also of the area in which they live. This area-level characterisation refers to both the physical characteristics of the focal area as well as its relative spatial configuration with its neighbouring regions.

Further, in order to proceed with the spatial analysis, urban spatial data is required, i.e., building footprints. Spatial data on DUAs are outdated or even non-existent, creating a gap in knowledge of the physical characteristics of deprivation. To generate such data, manual digitizing is the most conventional way of mapping, but it requires expert interpretation and field work, which is costly and time-consuming. Moreover, it falls short in terms of updatability, scalability, and transferability. Hence, a more efficient methodology is needed to address these challenges, while also responding to the UN's demand to update data for the Sustainable Development Goals (SDGs) (UN-Habitat, 2015). Clearly, there is a need to improve data availability, quality, consistency, timeliness, and disaggregation in DUAs (Kuffer et al., 2021).

Remote sensing (RS) has shown good potential in the extraction of urban physical features by providing spatial disaggregated and geo-located urban data (Camps-Valls, Tuia, Xiang-Zhu, & Reichstein, 2021; Jochem & Tatem, 2021). Additionally, it can provide up-to-date data due to its constant monitoring of the Earth's surface. Spatial Analysis (SA) applications of RS data have been explored in recent years to provide operational support for the characterisation and measurement of urban forms (Chen, Leu, & Wang, 2019). Specifically, it is a useful source of data to characterise urban morphology (Ibrahim, Haworth, & Cheng, 2021) and apply it to urban planning (Zhong et al., 2010).

Deep learning (DL) algorithms have obtained remarkable results in many computers vision tasks (Lecun, Bengio, & Hinton, 2015) and are recently being applied successfully to the processing of RS data for generating building maps (Camps-Valls et al., 2021; Wu et al., 2018), offering better performance for object detection and image segmentation compared to traditional algorithms (Li, Wang, Wang, & Lu, 2018). Specifically, convolutional neural networks (CNN) and deep neural networks (DNN), algorithms improve considerably the performance of semantic segmentation (Krizhevsky, Sutskever, Hinton, & G., 2015; Simonyan & Zisserman, 2015). The conclusions in recent literature on the best performing deep learning architecture for building footprint extraction from VHR imagery suggest that U-Net is one of the best options (Ayala, Aranda, & Galar, 2021; Rastogi, Bodani, & Sharma, 2020; Peng, Zhang, & Guan, 2019; Li, Wang, Zhang, & Zhang, 2019; Xing et al., 2019). Furthermore, U-Net is used by companies such as Microsoft and Google for mapping building footprints (Sirko et al., 2021; Yang, 2018). However, as both data sets have not been trained with enough buildings from Deprived Urban Areas (DUAs), they perform poorly and inconsistently within DUAs.

Mapping applications of DUAs through RS-based methods have increased since the availability of more very-high-resolution sensors (Kuffer, Pfeffer, & Sliuzas, 2016; Mason & Fraser, 1998). In the past decade, many studies highlighted the capacity of RS to map DUAs (Kuffer et al., 2020; Ranguelova et al., 2018). Several studies stressed the utility of contextual features for DUA mapping (Engstrom, Harrison, Mann, & Fletcher, 2019). Such aspects capture the built-up morphology. Recently, there has been a growing body of literature that maps DUAs using DL, i.e., by using CNNs (Ajami, Kuffer, Persello, & Pfeffer, 2019; Wang, Kuffer, Roy, & Pfeffer, 2019) or FCNs (Wurm, Stark, Zhu, Weigand, & Taubenböck, 2019).

Several studies have detected DUAs areas mainly at the city scale, focusing on the boundary of the settlements (Kohli, Sliuzas, & Stein, 2016; Kuffer & Barros, 2011; Wurm & Taubenböck, 2018). For this, an automatic global process remains a challenge due to their variability in types and definitions between cities (Duque, Patino, Ruiz, & Pardo-Pascual, 2015; Kuffer, Barros, & Sliuzas, 2014; Sliuzas & Kuffer, 2008a, 2008b) and due to their often-rapid development processes (Liu, Kuffer, & Persello, 2019). Semantic segmentation of buildings in DUAs from RS-based methods is still challenging due to the urban complexity (i.e., variety of physiognomy and materiality of the roofs of buildings) (Fig. 1). CNN building segmentation in a DUAs was seen for the first time in Guangzhou City, Southern China (Pan, Xu, Guo, Hu, & Wang, 2020) and in Ahmedabad city in Gujarat, India. The U-net architecture was adopted for building segmentation, showing robust performance. However, the spatial urban characteristics found in Nairobi City are different from those in Guangzhou and in Ahmedabad, where the roofings are more regular, with similar materials and the space between buildings is larger.

Cities are complex systems where diverse domains interact in a static physical structure (White, Engelen, & Uljee, 2015). The urban form facilitates the social, economic, and cultural life of the city, and when a design is inadequate, processes are hindered (e.g., mobility). Therefore, characterizing and mapping the urban form is essential to understand the flow of different domains (Wurm & Taubenböck, 2018), relating spatiality to urban behaviour, and guiding evidence-based planning accordingly (Simon, 1962). The physical structure of cities can be further understood as being characterised by quantifiable elements, also known as morphological features. This urban morphology can be defined by its fundamental elements, i.e., buildings and open spaces, plots, and streets (Mumford, 1961), and it can be quantified by a set of geometric measures – indicators – characterizing the different urban spaces. The morphological characterisation helps in capturing the high degree of variation observed across urban forms. Meanwhile, this variation is not just limited to the differences between non-deprived and deprived urban areas (DUAs), but in fact as Davis (2006) points out there exist intrinsic differences between and within DUAs themselves Davis (2006). This morphological characterisation can be an expression of inequality and socio-economic disparities visible from both facets: from

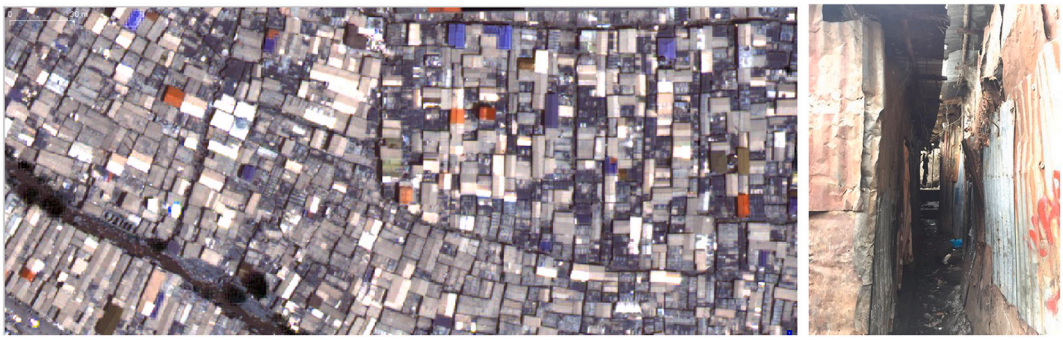

Fig. 1. WorldView-3 imagery (resolution: 0.3 m). A dense urban slum in Nairobi. Variable building sizes and compact urban form. Right: Ground photo taken by the author. The roofs overlap at different levels and the street below cannot be detected from the satellite image.

the street and from space (Naik, Raskar, & Hidalgo, 2016; Sliuzas & Kuffer, 2008a) (Fig. 2).

Although DUAs have existed in cities from their beginning and some governments acknowledge their existence (Davis, 2006), deprivation research has greatly increased following the third United Nations Conference on Human Settlements (Habitat III) in 2015, calling for an SDG assessment. Since then, the scientific literature has become increasingly interested in studying deprived urban characterisation (in response to SDG 11). Urban spatial deprivation has often been simplified by mapping socio-economic data aggregated to administrative units, as poverty has predominantly been characterised as household-level deprivation based on census data (Schirmer, van Eggermond, & Axhausen, 2014). Meanwhile, the area-level characterisation in DUAs is mostly ignored (Abascal et al., 2021; Baud, Kuffer, Pfeffer, Sliuzas, & Karuppannan, 2010; Kuffer et al., 2014; Taubenböck et al., 2018; Thomson et al., 2020). Spatial characteristics of deprivation are still unknown and there is even less global agreement to formulate deprivation through morphological variables.

Few studies have acknowledged the diversity within DUAs (e.g., Graesser et al., 2012; Krishna, Sriram, & Prakash, 2014, Kuffer, Pfeffer, Sliuzas, Baud, & van Maarseveen, 2017; Wurm & Taubenböck, 2018), and the categorization has mostly been based on statistical methodologies (e.g., PCA). Nonetheless, local expertise (i.e., link with local urban policies) and ground knowledge (i.e., DUA inhabitant's insights) have been important in interpreting and categorising deprivation (Joshi, Sen, & Hobson, 2002). The physical characterisation in quantifying deprivation have been conducted mostly based on household-level characteristics, such as building conditions which are defined by building size or building material, and primarily extracted from census or survey data (Anurogo, Lubis, Pamungkas, Hartono, & Ibrahim, 2017). Deprivation conceptualisation through urban physical characteristics is still largely under-researched, without an agreement on the indicators conceptualisation nor the methods of measurement. On the other hand, the physical conditions of the area that are analysed by RS-based methods (Kohli, Sliuzas, Kerle and Stein, 2012; Taubenböck et al., 2018) are mainly focused on isolated features such as buildings and open spaces within a settlement. Both the relative location of the area within the larger urban spatial configuration as well as the subtler aspects of the building orientation are found missing from the literature. Therefore, there is a need for an integrated approach that situates these isolated features within a larger spatial configuration and considers the orientation within the settlement pattern.

This paper aims to investigate the following research question: Can deep learning be used to characterise degrees of deprivation based on the morphology of DUAs in LMICs? To this end, the following specific objectives are pursued:

1. To generate a reference dataset of building footprints in DUAs in LMICs through participative community-based crowdsourcing
2. To employ deep learning for the automatic generation of building footprints

Fig. 2. Top row: RS imagery (Google Earth) of Nairobi city. Bottom row: 3 images from ground level; an example of a deprived area image taken by the author (left: red frame), a low-cost housing area (centre: blue frame) and a middle-class housing area (right: yellow frame), the latter two from Google Street View. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. To perform a morphological analysis based on the building footprints from deep learning (predicted buildings) and on the participative crowdsourcing reference dataset (digitised buildings) and link its output to deprivation levels

We claim that morphological characteristics can be captured by RS and be linked with deprivation. The paper is organised as follows: Section 2 covers the Study Area and Data used, Section 3 the Methodology, Section 4 the Results, Section 5 the Discussion and Section 6 the Conclusion.

2. Study area and data

2.1. Study area

This study is conducted in Nairobi City, Kenya. Nairobi's first settlement dates to the year 1899 and was set up as a colonial railway settlement (Morgan, 1967). The political responses to the basic right to housing has been inadequate since the colonial era, and the continued existence and growth of DUAs constitute a real challenge to urban life in Nairobi (Van Zwanenberg, 1972). Nowadays, more than half of its population live in deprived urban areas, which cover less than 6% of the total city area (APHRC, 2012).

The administrative area of Nairobi city covers 695 km^2, while the total area of the urban deprived regions within the city is 7,75 km^2 (APHRC, 2012). The administrative boundary remains the same as in the year 1963, just after the Kenyan independence. It is now outdated, as the city has expanded outwards, and it is in this periphery that new DUAs are emerging today. Based on the data availability and contacts with community-based groups, the DUAs analysed in this study are limited to ones that are within the Nairobi city administrative area. Nairobi city is divided into sub-counties (11), divisions (29), locations (72) and the lower disaggregated unit are the sublocations (147) (KNBS, 2019). Despite these fine-scale subdivisions (e.g., smaller sublocation size is 0.07 km^2), not all DUAs are delimited within an administrative boundary, and therefore cannot be adequately analysed through a census analysis.

From its earliest times, spatial patterns in Nairobi reflected divisions in terms of social class, mostly related to race due to colonial planning. This segregation was between the Central Business District (CBD) and European, Asian, and African residential areas (Morgan, 1967). Today, this structure is reflected not so much in terms of race, but instead in terms of spatial and economic deprivation (K'Akumu & Olima, 2007). The wealthiest people live in Nairobi West, which is characterised by its greener areas and lower built-up density. On the other hand, low- and middle-income groups dominate in the eastern locations (Fig. 3). However, within the areas populated by the low-income groups itself, one observes differences in terms of economic deprivation and the corresponding physical characteristics.

The DUAs analysed are shown in Fig. 3, marked with white boxes, namely Korogocho, Mathare, Kariobangi, Waruku, Pumwani, Soweto, Kibera, Mukuru, and Mukuru Kwa Njenga.

2.2. Datasets

2.2.1. Satellite imagery

VHR images were acquired by the Worldview-3 (WV-3) satellite in 2019. They cover Nairobi City County almost entirely (excluding the Nairobi National Park) while extending to part of the peripheric urbanised areas that fall outside of the Nairobi administrative boundary (blue polygon in Fig. 3). For this study, only the DUAs area was used with total coverage of 7 km2 (more details in Table 2). The WV-3 data consist of eight multispectral bands (MS) (1.20-m resolution) and a panchromatic band (0.30-m resolution). The WV-3 MS bands contain information across the visible and near-infrared spectrum (coastal, blue, green, yellow, red, red edge, NIR 1, and NIR 2 bands). Table 1 shows the satellite imagery specifications (DigitalGlobe, 2014).

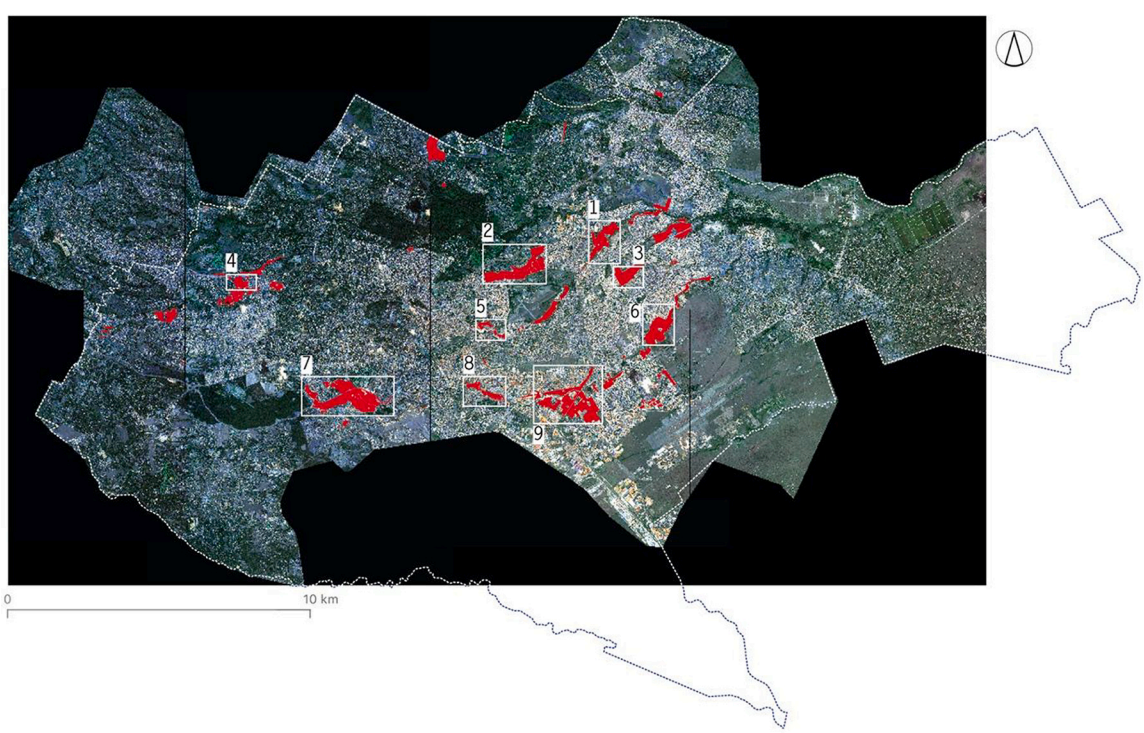

Fig. 3. WV-3 image shown as RGB colour composite of the City of Nairobi. Blue polygon represents the Nairobi administrative boundary. Red polygons indicate "slum areas" provided by Spatial Collective. The ones inside the white boxes are used in this research. 1. Korogocho; 2. Mathare; 3. Kariobangi; 4. Waruku; 5. Pumwani; 6. Soweto; 7. Kibera; 8. Mukuru; 9. Mukuru Kwa Njenga. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
WV-3 images specifications.

Bands	Spatial resolution	Geometric processing level	Radiometric processing level	Date
Panchromatic	0.30 m	Map Ready Ortho	Automatic atmospheric correction	Mosaic of 2 images (13/01/2019 and 01/02/2019)
8 multispectral bands	1.20 m	Map Ready Ortho	Automatic atmospheric correction	Mosaic of 2 images (13/01/2019 and 01/02/2019)

2.2.2. *Reference data*

The reference data consist of manually delineated buildings footprints (digitised buildings). Three entire settlements were selected, namely Mathare, Korogocho and Mukuru, as well as parts of other settlements (Kariobangi, Waruku, Pumwani, Soweto, Kibera, Mukuru Kwa Njenga). Different characteristics guided their selection i.e., diversity in size and shape of buildings (as small-sized buildings are commonly found in DUAs), diversity in roofing materials, diversity in urban patterns, diversity in open space forms, and diversity of street widths.

First, we obtained a building footprint layer generated by the community of Mathare through participatory mapping coordinated by the Nairobi-based Spatial Collective (SC) company. A total of 4410 buildings in Mathare North (locally known as Mlango Kubwa) were digitised using a Google Satellite image of 2019 as background. The original dataset can be found in SC GitHub repository (SCollective, 2020). Then, we inspected the building footprints and corrected the digitizing errors. Subsequently, we extended this dataset to the other DUAs through photo interpretation (over the WV3 reference image from 2019) and digitised 25,000 additional buildings (Zenodo, 2022). During this process, community members carried out targeted ground visits to validate a number of digitised footprints for which photo interpretation was uncertain.

3. Methodology

The methodology section is structured in two steps (Fig. 4). The first

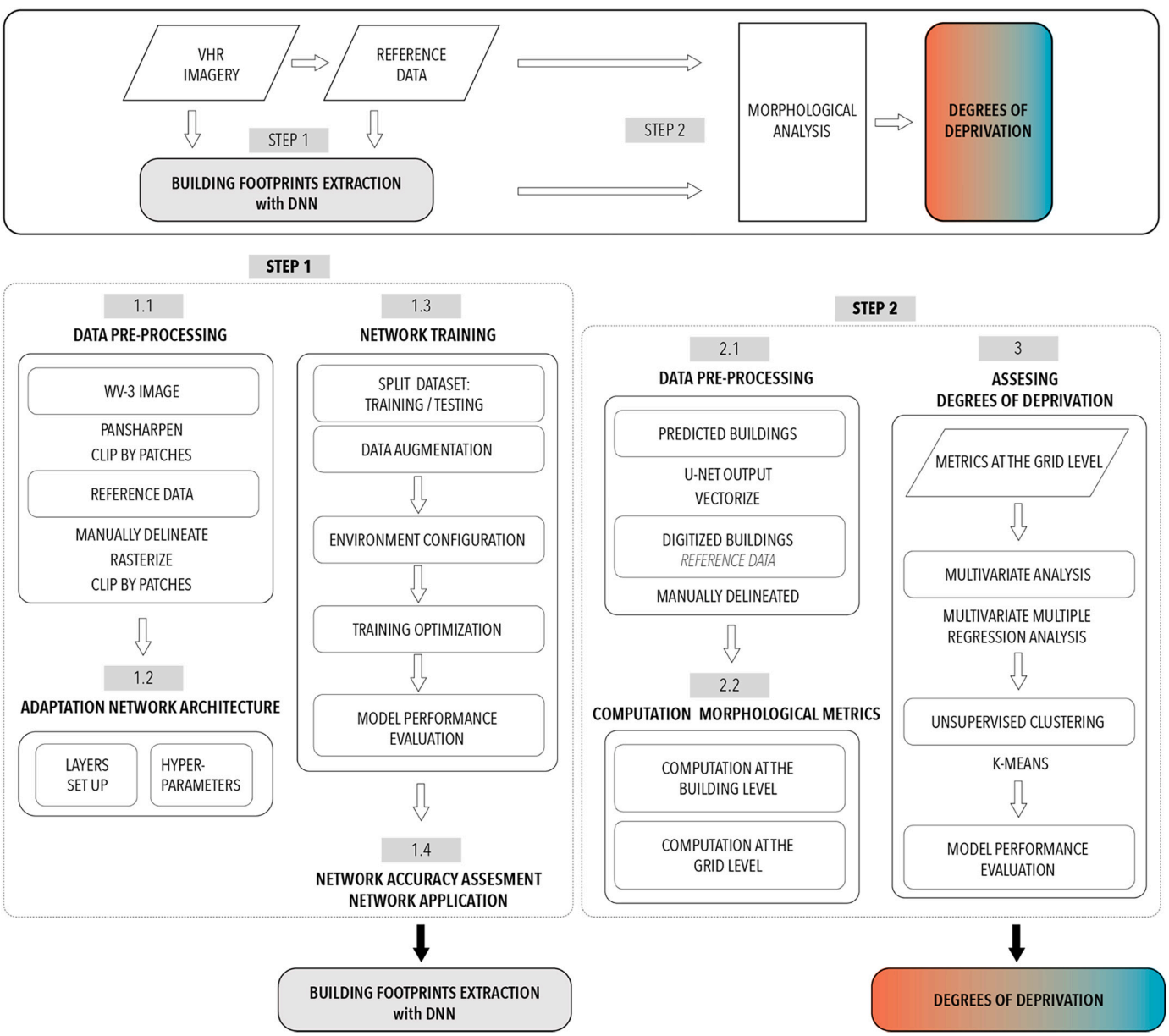

Fig. 4. Overview of the main steps of the methodology.

step introduces the extraction of the building footprints (Section 3.1), and the second step details the computation of the morphological metrics and the analysis of these metrics in link with deprivation levels (Section 3.2 and Section 3.3).

3.1. Building footprints extraction with DNN

The extraction of the building footprints was performed with a DNN and is structured as follows: (1) Data Pre-processing; (2) Adaptation of the Network Architecture; (3) Network training; (4) Validation; (5) Application. The complete procedure is shown in Fig. 4 (step 1).

3.1.1. Data pre-processing

The reference dataset needs to be split while training the DNN model in training (data used to fit the model), validation (parameter optimization of the model) and testing datasets (unbiased evaluation of final model (Kuhn & Johnson, 2013). As shown in Table 1, data from Mathare, Korogocho and Mukuru was used for training and validation of the deep learning model, whereas data from Mathare Central, Kariobangi, Waruku, Pumwani, Soweto, Kibera and Mukuru Kwa Njenga provided unseen data for testing. This choice was motivated by the need to test the generalisation ability of the deep learning model across datasets exhibiting spatial diversity.

The Coastal band was discarded due to atmospheric noise. The WV-3 MS bands were pansharpened with Gram-Schmidt Average Neighbourhood method based on the nearest neighbourhood diffusion pan sharpening algorithm (NNDiffuse) (Sun, Chen, & Messinger, 2014). Next, the images were clipped in a grid of 512*512 px and split into the datasets (Table 2).

The reference data were manually delineated by community members and by authors in the QGIS software with WGS 84/ UTM zone 37S as coordinate reference system. The reference data was rasterised with the algorithm Rasterize of QGIS 3.12, using 0.30 cm as pixel resolution. Afterwards, the raster was clipped in a grid of 512*512 pixels and split similarly to the satellite images.

3.1.2. Adaptation of the deep learning architecture

The U-Net architecture proposed in Ronneberger, Fischer, and Brox (2015) is adapted in this work. The network has a characteristic encoding-decoding structure. As shown in Fig. 5, feature maps from one encoding layer are concatenated to the corresponding decoding layer using skip connections. In the contracting branches, downsampling is achieved using maxpooling with a size of 2 × 2. On the other hand, upsampling is achieved through bilinear upsampling with a factor of 2 in the expanding branches.

3.1.3. Network model training

In total, 276 patches comprising seven bands and dimensions of 512 × 512 pixels were generated with an overlap of 64 pixels (px) (to cover unconcluded buildings after being clipped). The dataset was randomly partitioned into 75% training, i.e., 192 patches and 25% validation, i.e., 64 patches. To expand the training dataset, data augmentation was applied to the training patches. Geometric augmentation was done by rotating (45, 90, 135, 180, 270) and flipping (horizontal and vertical), allowing the network to learn invariance to such changes. This is an interesting approach not only because it provides more training data, but also because DUAs do not follow any spatial pattern, and all possible orientations can be simulated. During training, the categorical loss function is minimised (Bishop, 2006), and the optimisation done using Adam optimizer (Kingma & Ba, 2015) with a learning rate of 0.0001. In case the validation accuracy fails to improve after 19 epochs, the training is stopped.

The U-Net was implemented based on the *Tensorflow (Tensorflow 2.0)* and *Keras* (Keras 2.2.4) deep learning frameworks with *Python* (Python 3.7.3) programming language, on a computer running on Linux 5.10, with Intel i7, 32GB RAMand Nvidia GeForce RTX 2070 (8GB). For further details, the code is available at Zenodo scientific repository (Zenodo, 2022).

Table 2
Split of the reference dataset.

Dataset	DUAs	Purpose	Area
1	Mathare North Mathare South Korogocho Mukuru	**Training** and **validation** dataset	6,50 km2
2	Mathare Central Kariobangi Waruku Pumwani Soweto Kibera Mukuru Kwa Njenga	Testing dataset	0,5 km2

3.1.4. Accuracy assessment

Accuracy assessment of the U-net model performance for extracting building footprints in DUAs was evaluated by the confusion matrix (Table 3).

TP is known as True Positive, FP as False Positive, TN as True Negative and FN as False Negative. Recall and Precision scores were extracted from the confusion matrix represented by the ground truth with the predicted values.

$$Recall = \frac{TP}{TP + FN}, Precision = \frac{TP}{TP + FP} \tag{1}$$

The F1-score evaluation is derived from the confusion matrix as the harmonic mean of Precision and Recall. It can be interpreted as a weighted average of the precision and recall and varies from 1 (best) to 0 (worst). The relative contributions of precision and recall to the F1-score are equal.

$$F1 - score = 2*\frac{Precision*Recall}{Precision + Recall} \tag{2}$$

The Jaccard Index, also known as Intersection-over-Union (IoU) was calculated as an additional accuracy metric. IoU is the area of overlap between the prediction (A) and the reference (B) divided by the area of union between the prediction (A) and the reference (B). IoU tends to have lower accuracies than F1-score, as it penalises misclassification.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \tag{3}$$

3.2. Computing morphological metrics

Physical differences in urban patterns from a fine-scale morphological perspective were computed. Mathare Central, locally known as Huruma, was selected since it comprises heterogeneous morphologies (Fig. 6).

Urban form is assessed through morphological indicators related to the building form and the open space form (Mumford, 1961). A subset of metrics is derived from each indicator. Morphological metrics are usually based on clearly demarcated building outlines and are categorised according to 1) building form (size and internal irregularity) and 2) open space form (proximity and directionality). However, since U-Net building predictions can result in building clumps rather than individual buildings, especially where the urban area consists of densely clustered buildings constructed using a variety of hard-to-distinguish materials, we propose a series of equivalent metrics to be computed at a grid level for the predicted buildings, as shown in Table 4. (See Table 4.)

Different levels (scales) were used in the morphological analysis, as shown in Fig. 7. The digitised buildings metrics were computed at the building level and aggregated at the grid level. The predicted buildings

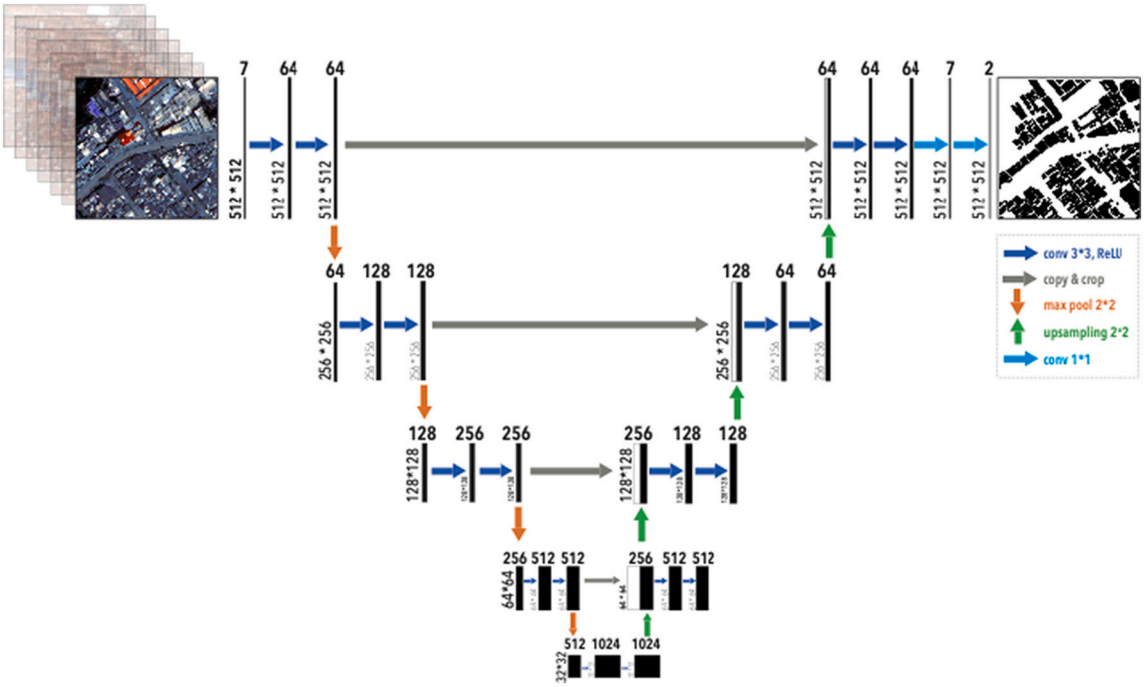

Fig. 5. An adaptation of U-net architecture (Ronneberger et al., 2015). Each black rectangle represents a multi-channel feature map. The number of channels is denoted at the top of each. The left label denotes the size of the patch. Coloured arrows represent the different operations.

Table 3
Confusion matrix.

Ground truth			
Prediction		Roofs	No roofs
	Building	TP	FP
	No building	FN	TN

metrics were computed at the grid level. At the grid level both building datasets were characterised and deprivation clustering was performed (section 3.3). There is no building footprint coverage in DUAs, so there are no settlement level metrics for the digitised buildings. The settlement level was used to show deprivation degrees results within the studied DUAs in Nairobi City.

3.2.1. *Computation at the building level – Digitised buildings*

The building level defines the attributes related to the physical characteristics of the building form as well as its proximities (e.g., size and arrangement). Several aspects were considered while performing the morphological analysis, reflected in four main indicators: size, inner irregularity, proximity, and directionality. Thirty-five metrics were extracted related to statistics from each indicator (Table 3).

Related to the building form only the shape characteristics were chosen displayed in building size and inner irregularity. The condition of the buildings (e.g., building materials) was not considered, since only the roofs are visible on monoscopic satellite images. Techniques such as 3D building models (Taubenböck & Kraff, 2014) were not implemented, as Google Street View is not available for most DUAs in the city of Nairobi. Seven metrics were extracted relating to the indicator building size. Among the variables related to the inner irregularity of the

Fig. 6. Mathare Central area (See the location of Mathare DUAs at the city scale on label 2 of Fig. 3).

Table 4
Metrics at the grid level. **Thirty-five** digitised building metrics. **Twenty-six** predicted building metrics.

Urban form	*Indicators*	*Metrics - digitised buildings*	*Metrics - predicted buildings*
Building form	Building Size	"area"	"area"
		"perimeter"	"perimeter"
		Length façade ("long side", "short side")	–
		Number vertices building ("num_vertices")	Number vertices clump ("num_vertices")
		"eccentricity"	–
		"elongation"	–
		Number of buildings ("num_build")	Number of clumps ("num_clumps")
	Inner Irregularity	Internal angle (inn_angle) min ("inn_angle_min")	–
		Inn_angle maximum ("inn_angle_max")	–
		Inn_angle mean ("inn_angle_mean")	–
		Inn_angle irregularity ("inn_irr_angle")	–
Open Space form	Proximity	Distance between centroids (dcc) min ("dcc_min")	–
		Dcc maximum ("dcc_max")	–
		Dcc mean ("dcc_mean")	–
		Dcc range ("dcc_range")	–
		Dcc median ("dcc_med")	–
		Dcc standard deviation (sd) ("dcc_sd")	–
		Distance between nearest vertices (dvv) minimum ("dvv_min")	Max length of DT long sides (max long side) minimum ("min_max_long_side")
		Dvv maximum ("dvv_max")	max long side maximum ("max_max_long_side"
		Dvv mean ("dvv_mean")	max long side mean ("mean_max_long_side")
		Dvv range ("dvv_range")	max long side range ("range_max_long_side")
		Dvv median ("dvv_med")	max long side median ("med_max_long_side")
		Dvv standard deviation ("dvv_sd")	max long side sd ("sd_max_long_side")
		Distance between vertex and nearest façade (dvl) minimum ("dvl_min")	Mean length of DT long sides (mean long side) minimum ("min_max_long_side")
		Dvl maximum ("dvl_max")	mean long side maximum ("max_max_long_side")
		Dvl mean ("dvl_mean")	mean long side mean ("mean_max_long_side")
		Dvl range ("dvl_range")	mean long side range ("range_max_long_side")
		Dvl median ("dvl_med")	mean long side median ("med_max_long_side")
		Dvl standard deviation ("dvl_sd")	mean long side sd ("sd_max_long_side")
	Directionality	Orientation index (oi) minimum ("oi_min")	–
		Orientation index maximum ("oi_max")	–
		Orientation index mean ("oi_mean")	Orientation of DT short side mean ("mean_oi_short")
		Orientation index range ("oi_range")	Orientation of DT short side range ("range_oi_short")
		Orientation index median ("oi_med")	Orientation of DT short side med ("med_oi_short")
		Orientation index standard deviation ("oi_sd")	Orientation of DT short side sd ("sd_oi_short")
		–	Angle between DT long sides (ang long sides) minimum ("min_ang_long_sides")
		–	Ang long sides maximum ("max_ang_long_sides")
		–	Ang long sides mean ("mean_ang_long_sides")
		–	Ang long sides range ("range_ang_long_sides")
		–	Ang long sides median ("med_ang_long_sides")
		–	Ang long sides sd ("sd_ang_long_sides")

building, the internal angle of each corner was measured, and diverse statistics were calculated (formula in Fig. 8). The more irregular the internal building angles are, the more irregular the shape of the building itself and the more informal the building is understood.

Another morphological feature that characterizes the physical structure is the open space form. To define it, proximity and directionality indicators were extracted. The proximity between buildings represents the open space width. Proximity approximation was computed by measuring the distance between the building of study and the four nearest buildings. As opposed to what was presented in other studies, our metrics were calculated with the four nearest buildings, instead of using ten neighbouring buildings (Taubenböck & Kraff, 2014) as the building surrounding open space captures the width of the street. The four buildings selected were the closest centroids from each quadrant of the Cartesian axes (see Fig. 8). Different metrics were performed to define the best approximation: distances between centroids (Taubenböck et al., 2018), named: dcc; distance between the nearest vertices, named: dvv; and the minimum distance between the vertex of the study building and the orthogonal distance to the nearest facade of the neighbouring building, named: dvl. The directionality of the open spaces is defined by the building orientation index, oi (formula 6 in Fig. 8), as a proxy for measuring the alignment between buildings (Taubenböck et al., 2018; Venerandi, Quattrone, & Capra, 2018). To capture the index, instead of using pairs or buildings (Taubenböck et al., 2018), the four nearest buildings were used. Different statistics were performed. DUAs are characterised by being unplanned, which is reflected by their irregular physical structure.

3.2.2. *Computation at the grid level – Digitised and predicted buildings*

The grid cell level defines the attributes related to the physical characteristics of the urban pattern as the aggregation of each individual building and the open space disposition. The grid level was selected, as compared to blocks, as there are often no administrative boundaries or recognised roads available from which blocks can be extracted in DUAs (Grippa et al., 2018). The selected grid unit size is 75 m × 75 m as it permits optimal granularity representing a high variety of urban form, while protecting privacy. For example, the total number of grids was 52 within Mathare. At this level, the metrics can be computed from both datasets (Table 3), i.e., in the manually delineated set (digitised building footprints), by aggregating the building level metrics to the grid level; and from the U-net output set (predicted building footprints) by computing metrics and deriving statistics within the grid.

To characterize the grid level from the predicted footprints, certain metrics were calculated directly from the building footprints, while others were calculated from Delaunay Triangulation (DT) in the open space, using the building footprint corners as vertices. DT was produced in QGIS 3.12. (Fig. 9). Triangles from DT are conformed by two long

Table 4
MMRA between metrics of both datasets. In blue digitised footprint metrics. In pink predicted footprint metrics. $N = 52$ grid cells.

MMRA	sum_ area	mean_ area	sd_ area	short _mean	short _sd	mean_ num_ vert	num_ build	mean_ inn_irr_ ang	max_ inn_irr_ang	mean_ dvl_ mean	sd_ dvl_mean	mean_ oi_mean	sd_ oi_mean
Adjusted R-squared	0,87	0,07	0,14	0,30	0,27	0,19	−0,14	0,033	0,41	0,57	0,27	0,78	0,17
Intercept	−1230	640,20*	722,60·	9,24	9,55*	11,16 ·	123,8	0,34	1,57	25,97	49,93*	−0,12	−0,12
area	0,61**	−0,0059	0,0011	−0,00004	−0,0024	−0,00025	0,029	−0,000015	−0,000087	−0,0022*	−0,0031*	−0,00002	0,0000012
area-neig	−0,57 ·	0,014	0,0085	0,0009 ·	−0,00023	−0,00036	0,081	−0,000017	−0,00013	−0,0023	−0,0032 ·	−0,000007	−0,000014
num_clumps	−20,05	−0,76	−0,61	0,02	0,016	−0,036	0,42	−0,0044 ·	−0,0088	−0,048	−0,14 ·	0,00066	−0,00096
num_clumps-neig	57,11*	1,14	0,96	−0,017	−0,0029	0,053	−2,07	0,0031	0,0049	−0,049	−0,073	−0,0031	0,00026
num_vert	−0,09	−0,0089	−0,004	−0,00026	−0,0003*	−0,00015	0,011	−0,000009	0,000064	−0,00037	−0,00049	−0,00002	0,0000055
num_vert-neig	0,17	0,0009	−0,003	0,00035	0,00018	−0,00014	0,012	0,000041	−0,00018	0,00096	0,00062	0,00001	0,00001
max_mean_long _sides	−0,36	0,34	0,61	−0,0011	0,000016	0,0048	0,071	−0,00018	−0,0053	−0,049	−0,054	0,0008	0,000095
max_mean_long_sides-neig	13,5	−0,88	−1,13	−0,039 *	−0,0025	0,032 ·	−2,54	0,00097	0,012*	0,012	0,041	−0,0012	0,0037
mean_mean_long_sides	−143,5	−28,75	−23,73	−0,11	−0,54	−0,073	71,93	−0,07	−0,26	−0,55	−3,52 ·	0,013	0,028
mean_mean_long_sides-neig	820,5*	26,26	33,27	0,044	0,62	0,11	−90,62	0,053	0,18	1,57	3,41 ·	0,037	0,089
mean_oi_short	−2526	1369*	1888*	18,14	23,45*	16,72	1198	1,7	3,21	18,33	75	1	−0,054
mean_oi_short-neig	12,520	−434,4	−933,5	−15,03	4,33	26,52 ·	−759,1	1,11	1,38	−66,21	14,33	−1,19	−0,055
sd_oi_short	10,290	−2542*	−3315*	−35,11	−36,46*	−16,5	655,7	0,27	0,73	−39,12	−10,92	−0,48	0,71
sd_oi_short-neig	−851	481,1	1484	−3,09	1,13	−31,69	−214,6	−2,66	−4,64	52,77	−17,68	2,15	−0,48
mean_ang_long_sides	−28,37	2,14	3,12	0,056	0,036	−0,035	1,71	−0,0093*	−0,019	0,0094	−0,046	0,0075**	−0,0018
mean_ang_long_sides-neig	−26,53	−1,25	−5,37	0,094	0,049	0,032	−7,2	4,20	0,036 ·	−0,004	−0,022	−0,0044	−0,00073

Significant codes: '***': p-value <0.001; '**': 0.001 < p-value <0.01; '*': 0.01 < p-value <0.05; '·': 0.05 < p-value <0.1

sides (with a similar length) and a short side. DT mainly implies that the short side of the triangle links two vertices of the same building footprint, whereas the two long sides link vertices of distinct footprints. Except, rarely occurring, when a footprint is isolated between others, then the triangles do not always follow this placement (Fig. 9).

Metrics referring to the building form were extracted for the building size indicator (Table 3). Inner irregularity could not be simulated due to the pixel-based output that did not allow to measure the angles of the predicted buildings. Related to the building size the "area" and the "perimeter" of the predicted buildings were extracted, as well as the sum of all the predicted buildings ("num_clumps"), and the total number of predicted building corners ("num_vertices"). Related to the open space form, proximity and directionality indicators were defined by Delaunay Triangulation (DT) measures. Related to the proximity indicator, both DT long sides lengths were measured ("max_long_sides"; "mean_long_sides"). Intuitively, the distance between buildings is related to the width of the streets (Fig. 9). Related to the indicator directionality, the orientation of the short side of the DT was measured (oi_short = 1-(|angle-45|/45), expecting it can show differences in the orientation pattern, as that the direction of the street could estimate the direction of the buildings ("oi_short"). If angle = 0° OR 90° then oi_short = 0; if angle = 45° then oi_short = 1. The angles of the short side ("short_ang") were extracted as they simulate the orientation of the urban pattern. The angles between DT long sides were extracted as we observed that they vary depending on the width and direction of the open space ("mean_ang_long_sides").

The Pearson correlation coefficient between each pair of variables was calculated to define the independent variables that most define variability in the morphological characterisation. Thirty-five metrics were computed and highly correlated metrics, $-05 > r > 05$ were discarded. Weakly correlated metrics were aggregated to the grid cell level. Principal Component Analysis (PCA) was performed with the final weak correlated metrics to reduce dimensionality, not by creating new variables, as the metrics would lose urban interpretability, but to check whether all selected metrics had a high loading in the PCA test, and if not, to remove those variables and reduce the set of metrics.

Furthermore, to capture spatial patterns not only metrics within a grid cell were studied, for each grid cell, the difference between its value (j) and the mean of the eight neighbours'(i) was calculated. The Neighbour Index (Ineig) is based on the heterogeneity index from Taubenböck and Kraff (2014), capturing the similarity of the neighbours in those metrics.

$$Ineig = \left(\frac{\sum_{i=1}^{N} |Vi1 - Vj|}{N} \right) \tag{10}$$

where Ineig = Neighbourhood index; N = total number of neighbours studied; Vi = metric value of the neighbour grid; Vj = metric value of the studied grid.

3.3. Assessing degrees of deprivation at settlement level

Urban spatial deprivation was categorised based on the morphological variables extracted from grid level metrics (Fig. 10). Transferability, scalability, and replicability were aimed, so we wanted to obtain as an output of the model an interpretable degree of deprivation (based on predicted building metrics). Thus, we wanted to avoid the process of manually digitising building footprints to interpret the clusters and their morphological indicators but to interpret them directly from the model output metrics.

Firstly, a multivariate analysis was performed between both datasets to select among the metrics those that are highly representative in both sets. Because we have multiple responses, i.e., diverse predicted footprint metrics related to each indicator, tests for separate regression parameters are insufficient. It is important to establish whether our independent variables (ν) affect all the variables in the index (dependent

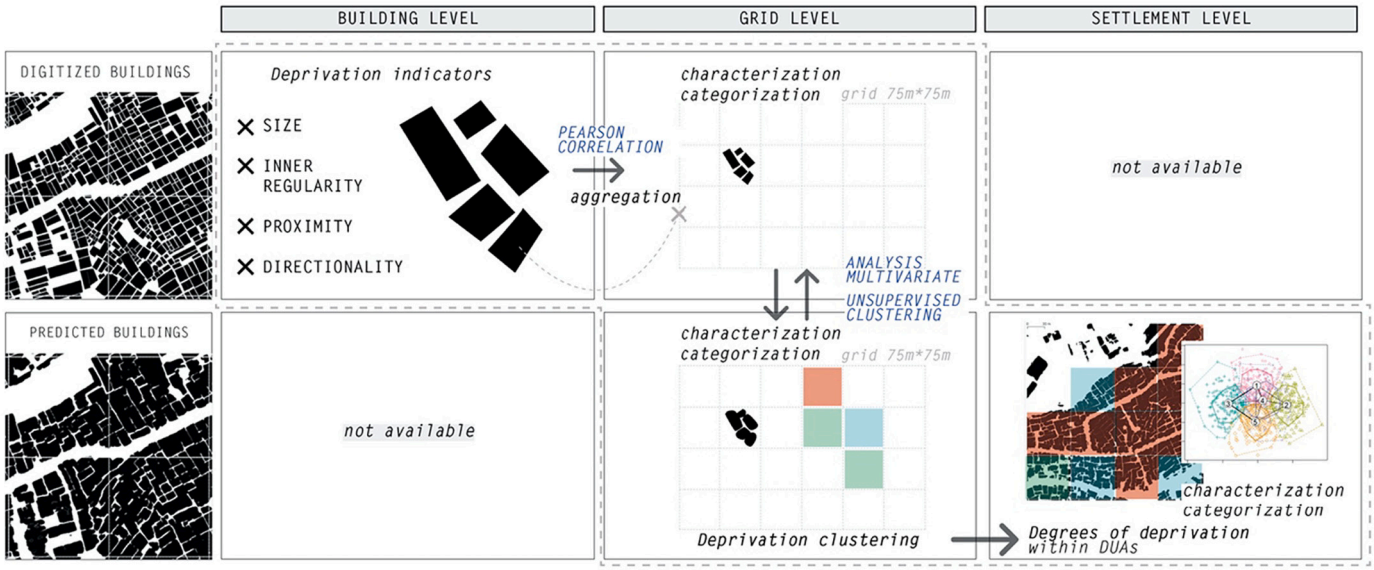

Fig. 7. Morphological analysis at different levels.

Indicators	Metrics	Formula	Geometry
Building Size	Area	$A = \frac{(L_1 \cdot h_1)}{2} + \frac{(L_2 \cdot h_2)}{2} + \ldots + \frac{(L_n \cdot h_n)}{2}$ (4)	A
Inner Irregularity	Intern Angle	$IIA = \frac{\max(\theta_1, \ldots, \theta_N) - \min(\theta_1, \ldots, \theta_N)}{\sum_{i=1}^{N} \frac{\theta_i}{N}}$ (5)	IIA = 0,5; IIA = 0
Directionallity	Orientation Index	$I_{01} = 1 - \lvert \Delta O - 45° \rvert \cdot \frac{1}{45} = 0$ (6)	AO = 0° Ioi = 0; AO = 90° Ioi = 0; 42°>AO > 3° Ioi = 0.6; 87°>AO > 48° Ioi = 0.6; AO = 45° Ioi = 1
Proximity	Distance between buildings	$dvlmin = \min(d_1, \ldots, d_5)$ (7); $dvlmean = \sum_{i=1}^{5} \frac{d_i}{5}$ (8); $dvlmed = \text{median}(d_1, \ldots, d_5)$ (9)	d5, d4, d1, d2

Fig. 8. Indicators, metrics, formulas, and geometry illustration at the building level.

variables, y) or only a few, i.e., whether all the digitised footprint metrics are explained by those in the model (predicted footprint metrics). For this purpose, the Multivariate Multiple Regression Analysis (MMRA) method was selected, to model the relationship between the metrics derived from the predicted footprints (independent variables) and the metrics derived from the digitised footprints (dependent variables).

$$y = \alpha_0 + \alpha_1{}^*\nu_1 + \alpha_2{}^*\nu_2 + \ldots + \alpha_n{}^*\nu_n \tag{11}$$

To achieve accurate results, both datasets assumed linearity, no multicollinearity, no outliers, similar spread across range and normality of residuals. MMRA was computed in the R statistical environment (R-3.6.1). We decided to keep the predicted footprints with a high level of significance ($p < 0.05$) and the digitised footprints with the adjusted R-squared value higher than 0.4, a threshold that captures approximately 23% of the standard deviation (Everit & Skrondal, 2010). The resulting MMRA metrics were used for cluster performance.

Secondly, the clustering process was carried out. We decided to go for an unsupervised method for both sets of metrics to propose a more objective and data-driven approach that can be interpreted in a more

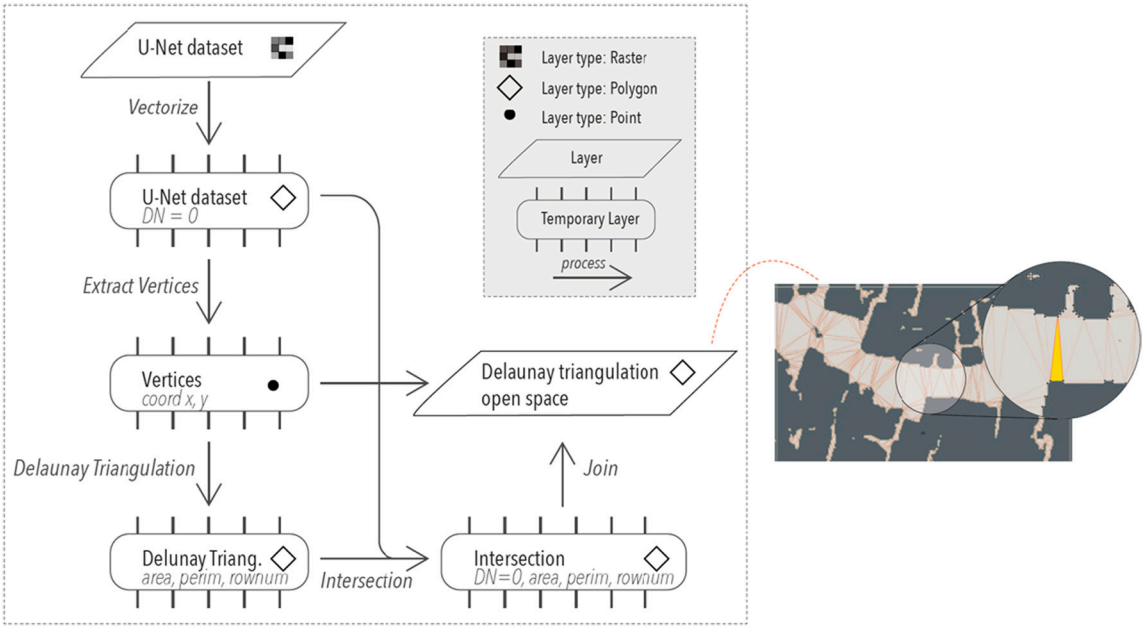

Fig. 9. On the left, DT procedure from U-net model output set at the grid cell. On the right, DT illustration.

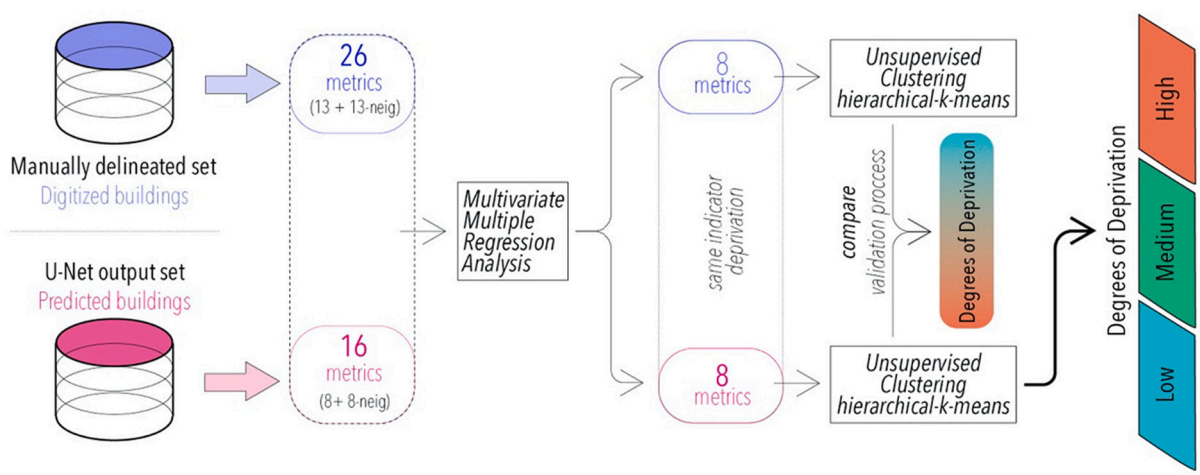

Fig. 10. Degrees of deprivation assessment procedure.

rigorous manner. For instance, creating references or labels (supervised methods) for such complex and largely unknown concepts as urban deprivation would bring a significant degree of uncertainty to our research.

Before performing clustering, the optimal number of clusters was determined. Several methods were tested to find the optimal number of clusters, using measures such as connectivity, the Dunn index, and the Silhouette index (Milligan & Cooper, 1985), using the Clvalid R package (Brock, Susmita, Pihur, & Datta, 2008). We proved varying the number of clusters from 2 to 6.

Different unsupervised clustering methods (k-means, k-medoids, hierarchical) were performed to guide the definition of degrees of deprivation. A hybrid method, hierarchical-k-means, was also performed as the post application of k-means optimizes and improves the initial partitioning generated by the hierarchical clustering method. All variables were scaled, to avoid the influence over the clustering of variables with greater value ranges.

The clustering method was selected based on the performance with the digitised footprint metrics, as it is the "reference clustering" and needs to be interpreted and validated. The same clustering method was performed with both sets of metrics. Clustering based on the digitised footprints metrics was used as a "reference clustering" for validating the clustering based on predicted footprints. The method selected was guided by the correspondence of each class and the levels of deprivation. Validation was carried out by consulting the current Kenya's building code legal document and through visual evaluation by local-based urban experts (i.e., urban planners and architects working in Nairobi).

The Building Code of Kenya (BCK) (Local Government, 1968) is the most up-to-date Kenyan legal code that delimits urban morphology. The BCK was analysed to interpret the morphological indicators and to relate them to degrees of deprivation. Concerning building form, only the building size indicator was assessed with the BCK, as no law is stipulated regarding the geometry of the building (i.e., inner irregularity). According to Act 72 (section 3 from the BCK) a "small house" is defined as a house with an area of less than 68 square metres (20,000 cu. feet). The higher the number of buildings per grid (num_build) in relation to the density (sum_area), the smaller the estimated area of each building, and the higher the degree of deprivation. As far as the open space form is concerned, only the Proximity indicator was taken as a reference for the BCK, as no law is stipulated regarding the Directionality of the urban

pattern. To define proximity, part 2 of the BCK was used, specifically Act 18, which stipulates 2.4 m (8 ft) as the minimum spacing between buildings. Act 17 mandates that each building must have a space in front of it of at least 6 m (20 ft). The distance between buildings was measured through the variables mean_dvl_mean and the sd_dvl_mean. They capture respectively the average and the sd per grid of the average of the distances of each individual building with the four nearest buildings. It is interpreted that the greater the distance between buildings, the less the area is deprived.

4. Results

4.1. Building footprint extraction with DNN

During the training process, the DNN model for extracting building footprints was evaluated in terms of accuracy-loss and the following values were obtained: Training Loss: 0.18; Training Accuracy: 0.92; Validation Loss: 0.24; Validation accuracy: 0.92.

The accuracy assessment of the model was also evaluated using the validation dataset (64 patches), at pixel level with the F1-score, and at area level with the Jaccard Index (IoU). The result from the **validation set** was **F1-score = 0.76** and **IoU = 0.61**. Besides the quantitative accuracy assessment, the results were also visually analysed, by comparing them to the reference data. The performance of the model on the validation dataset is shown in Fig. 11.

Subsequently, the testing dataset, composed of 80 patches (256 by 256 pixel) from nine different settlements, had an F **1-score = 0.84** and **IoU = 0.73**. The performance of the model in the testing dataset is shown in Fig. 12.

It was found that the predicted areas visually cover all the roofs. In certain areas, because of the high density of buildings and the barely existing space between them, the model had difficulties distinguishing the buildings individually. The WV3 pixel size in some areas is larger than the space between buildings. As illustrated in Fig. 1, overlapping of roofs is common in these areas. Therefore, in some cases it leads to a merging of buildings, showing one uniform polygon (clump) where several individual buildings should be observed.

4.2. Computing morphological metrics

Deep learning predictions face challenges in dense and closely packed urban areas where individual roofs are difficult to extract. In this type of urban setting, the objects that are extracted often represent building clumps (Fig. 13). For this reason, performing a morphological analysis at the building level would lead to inaccurate conclusions. Consequently, the morphological analysis needs to be conducted at an aggregated level, e.g., at the grid level.

Eight weakly correlated metrics from the digitised footprints extracted at the building level were selected (see Appendix). The selected eight metrics cover all four indicators. The Size Indicator was characterised by the area of each individual building ("area"), and the "perimeter" metric was discarded. The short side of the facade ("short") was chosen over "long" as some buildings share the same roof, and the short façade represents each individual house. The total number of corners ("num_vertices") was also selected. Eccentricity and elongation metrics ("ecc", "elong") were discarded due to high correlation with variables selected. The Inner irregularity Indicator was represented by the index of internal irregularity of their angles ("inn_irr_angle") calculated as shown in Fig. 10. Other variables within the indicator are correlated with it. In relation to Proximity Indicator, the min and mean distances between buildings were selected "dvl_min", "dvl_mean"). The minimum orthogonal distance between a vertex of the building and the closest facade of the neighbour (dvl) was selected over distance between centroids (dcc) and distances between vertex (dvv) as it better captures the real distance between buildings (i.e., street width). The minimum distance alludes to Act 18, and the mean to the Act 17 from the BCK. For

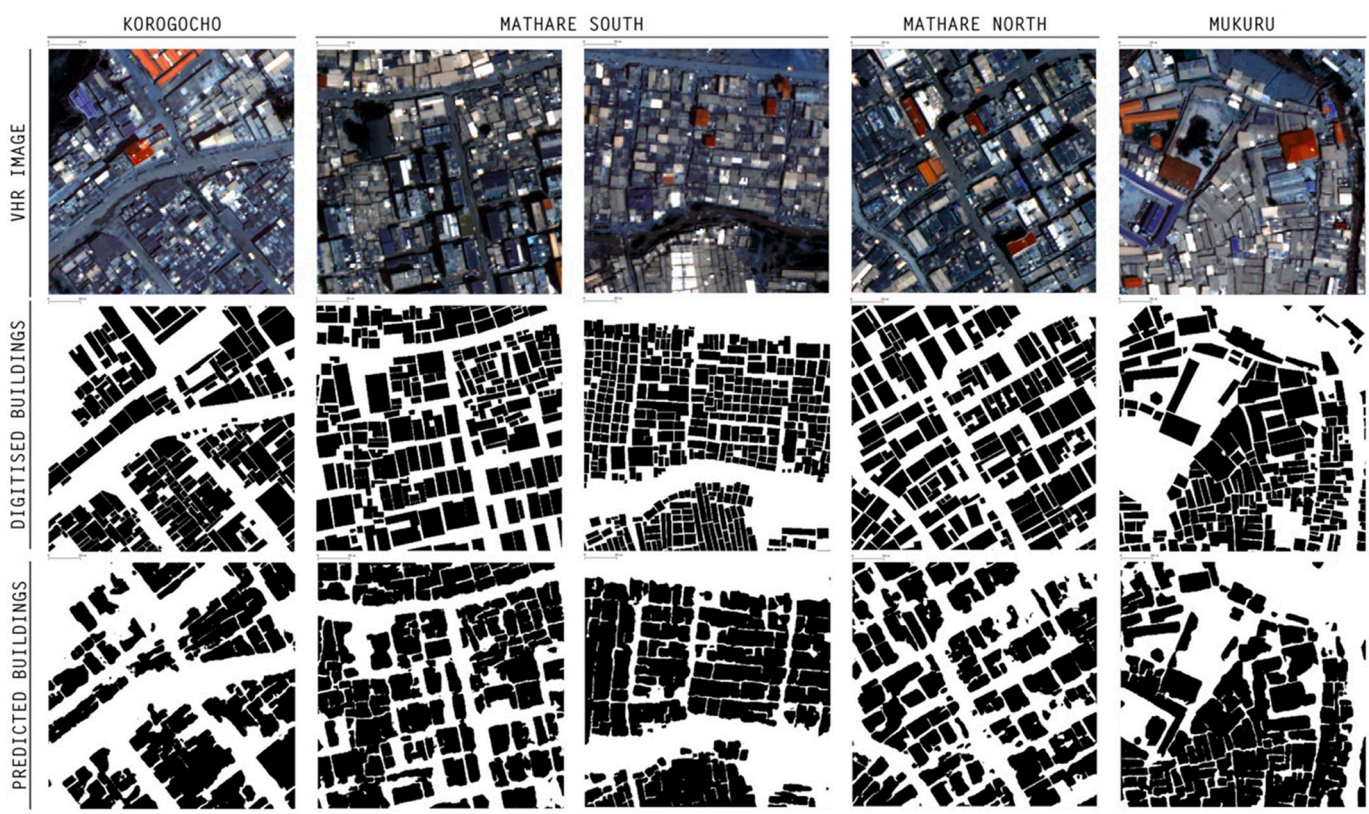

Fig. 11. Performance of building footprint extraction by U-net. Samples from all the DUAs. Top row: VHR Image, middle row: Digitised building footprint, bottom row: Predicted building footprint.

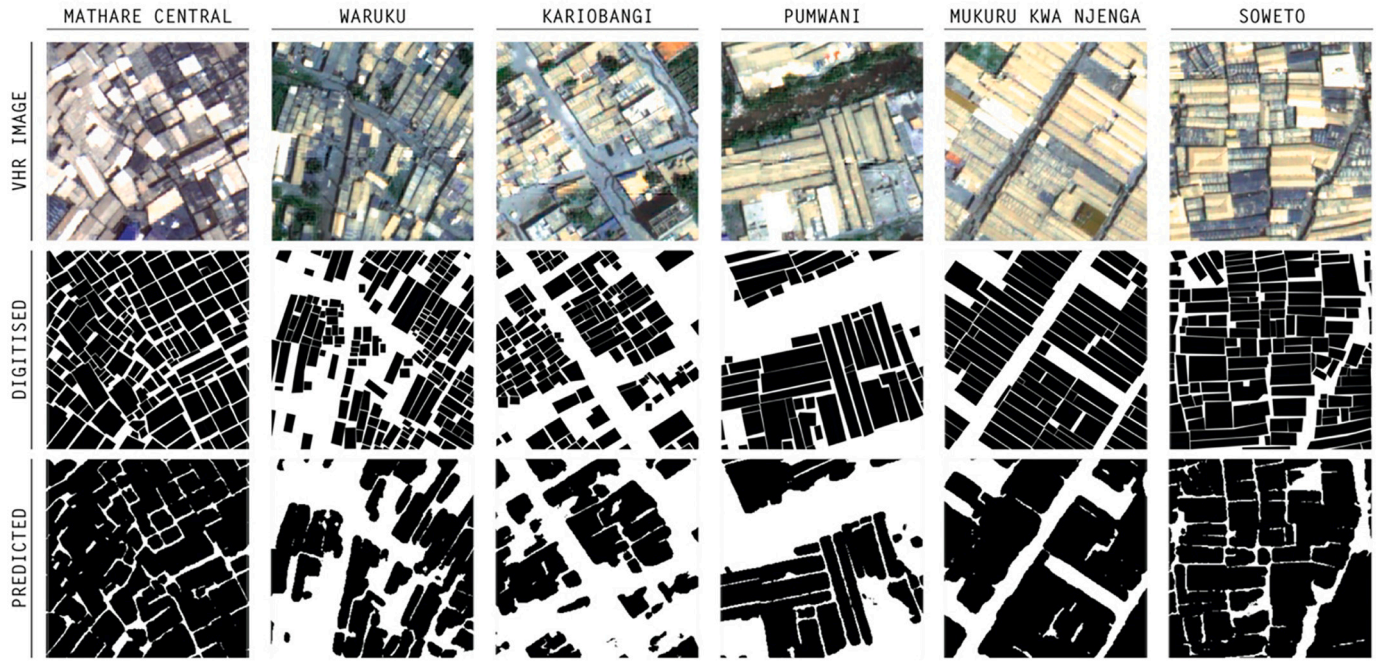

Fig. 12. Performance of building footprint by U-net using an independent testing dataset.
Samples from 6 deprived areas were tested. Top row: VHR Image, middle row: Digitised building footprint, bottom row: Predicted building footprint.

Fig. 13. Manually delineated set (left) and U-net output set (right) with centroids in each polygon. The manually delineated set consists of 4010 units, whereas the U-net output set contains 724 units.

the Directionality Indicator metrics "oi_mean" and "oi_sd" were chosen.

At the grid cell level, the eight metrics were aggregated by calculating their associated best-fit statistics (sum, min, max, mean, med, sd). The total number of buildings per grid ("num_build") was also calculated. Forty-nine was the total number of metrics aggregated to each grid cell. After computing Pearson Correlation coefficients, thirteen weakly correlated metrics were selected related to the four indicators. Size: "sum_area", "mean_area", "sd_area", "mean_short", "sd_short", "mean_num_vert" "num_build"; Inner Irregularity: "mean_inn_irr_ang", "max_inn_irr_ang". Proximity: "mean_dvl_mean", "mean_dvl_mean". Directionality: "mean_oi_mean", "sd_oi_mean". Then, for each weakly correlated metric, the neighbour index was also calculated, and aggregated to the grid. The Pearson Correlation Coefficient was calculated to measure the strength of correlation between the twenty-six metrics (Fig. 14).

Most grid cells show urban similarities with their surroundings i.e., neighbouring grid cells, which is reflected in the high correlation the metrics have with the neighbour (*metric*_neig). The more deprived the neighbours are, the more deprived the studied grid is. To capture spatial patterns, even if there were some metrics correlated with their neighbours, both metrics were kept to proceed with the deprivation clustering. Twenty-six total metrics were used for the deprivation clustering.

The predicted buildings metrics were extracted directly from each grid. Twenty-six was the total number of metrics aggregated to each grid cell. After computing the Pearson Correlation Coefficient, eight metrics were selected related to three main indicators. The final metrics were chosen as most alike to those of the digitised footprint metrics, i.e., the more interpretable. From the Size Indicator, metrics "area", "num_clumps", "num_vert" were kept. Proximity was defined by "max_mean_long_sides", "mean_mean_long_side" metrics. Both metrics were chosen, even if they present high correlation (0.75) as some grids show higher variability in their open space (e.g., very dense areas in the DUA edges are close to private, and not accessible, open spaces) and the "max_mean_long_sides" can detect this variability. The metric "mean_mean_long_side" also shows high correlation (−0.95) with the "area". As the built-up area increases, the proportion of open space is reduced, which is distributed and laid out in very narrow streets. Both metrics were kept as they define different deprivation indicators. The Directionality indicator was defined by "mean_ang_long_sides", "mean_oi_short", "sd_oi_short", metrics. The interpretability of these metrics in relation to the reference metrics is uncertain, so it was decided to retain them even though there is a high correlation. After performing PCA all metrics were kept since their loadings in at least one of the principal components obtained the largest absolute value. Finally, for each

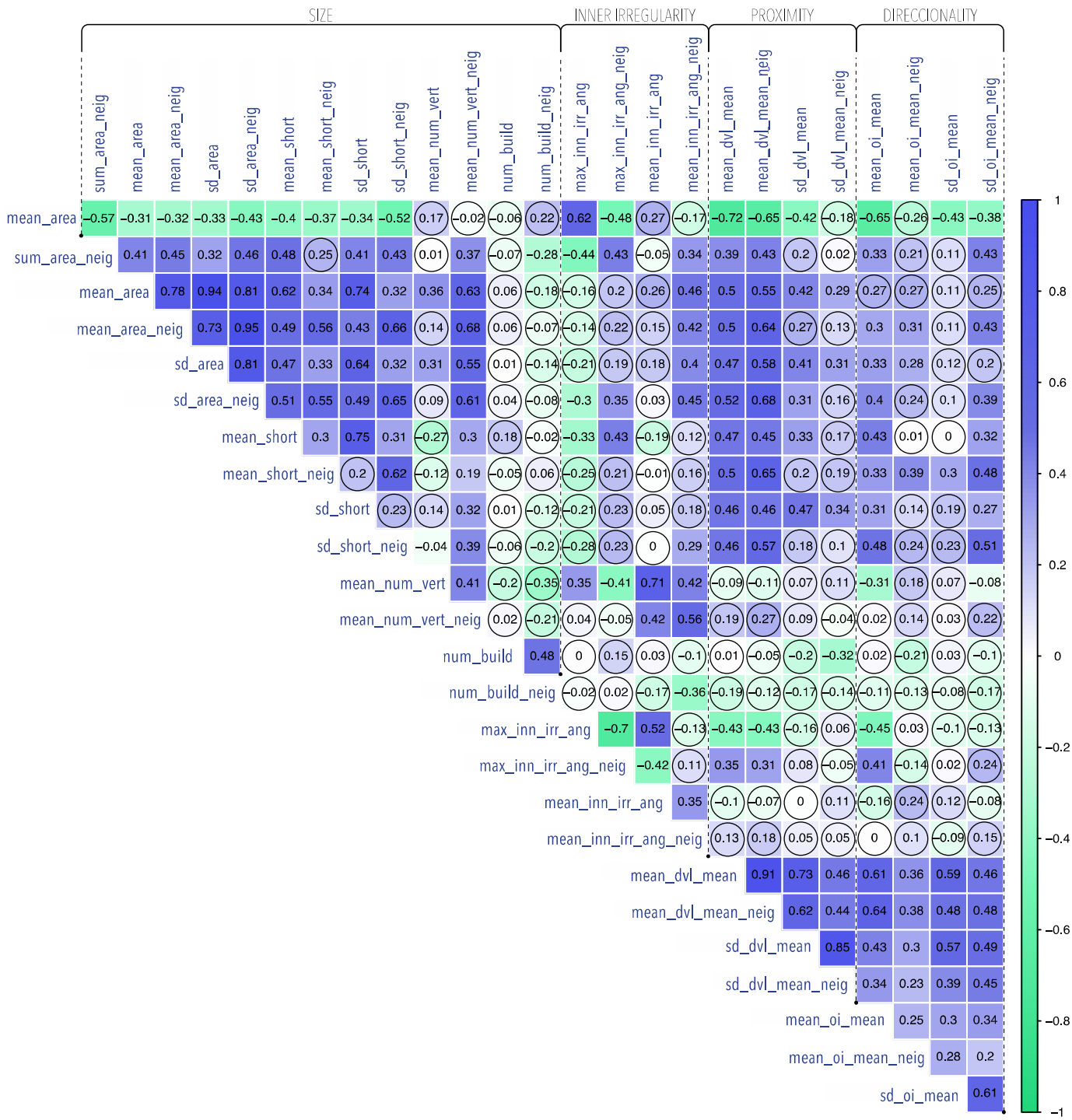

Fig. 14. Pearson correlation coefficients for **twenty-six** digitised buildings metrics calculated at the grid level including the neighbouring metrics. Values without circle refer to-significant *p*-values, $p < 0.05$.

variable the neighbour index was applied, and aggregated to the grid. The Pearson Correlation Coefficient was again calculated for the Sixteen metrics (Fig. 15).

4.3. Assessing degrees of deprivation from metrics

Table 4 shows the MMRA results of the digitised footprint metrics using the predicted footprint metrics as explanatory variables. For further analysis, we retained only statistically significant values ($p < 0.05$) with an adjusted R-squared value higher than 0.4. These values operate as a threshold that captures approximately 23% of the standard deviation explained by the digitised footprint metrics. Low values of the adjusted R-squared (as in num_build) indicate that the model does not capture the complexity of the relationships.

Since the goal was to capture deprivation from space (satellite imagery), **the interpretability** of the predicted metrics was of key importance. Therefore, we captured the relationships between the metrics of both sets that produce the same deprivation indicator, and we discarded the metrics that were not significant for modelling the degrees of deprivation.

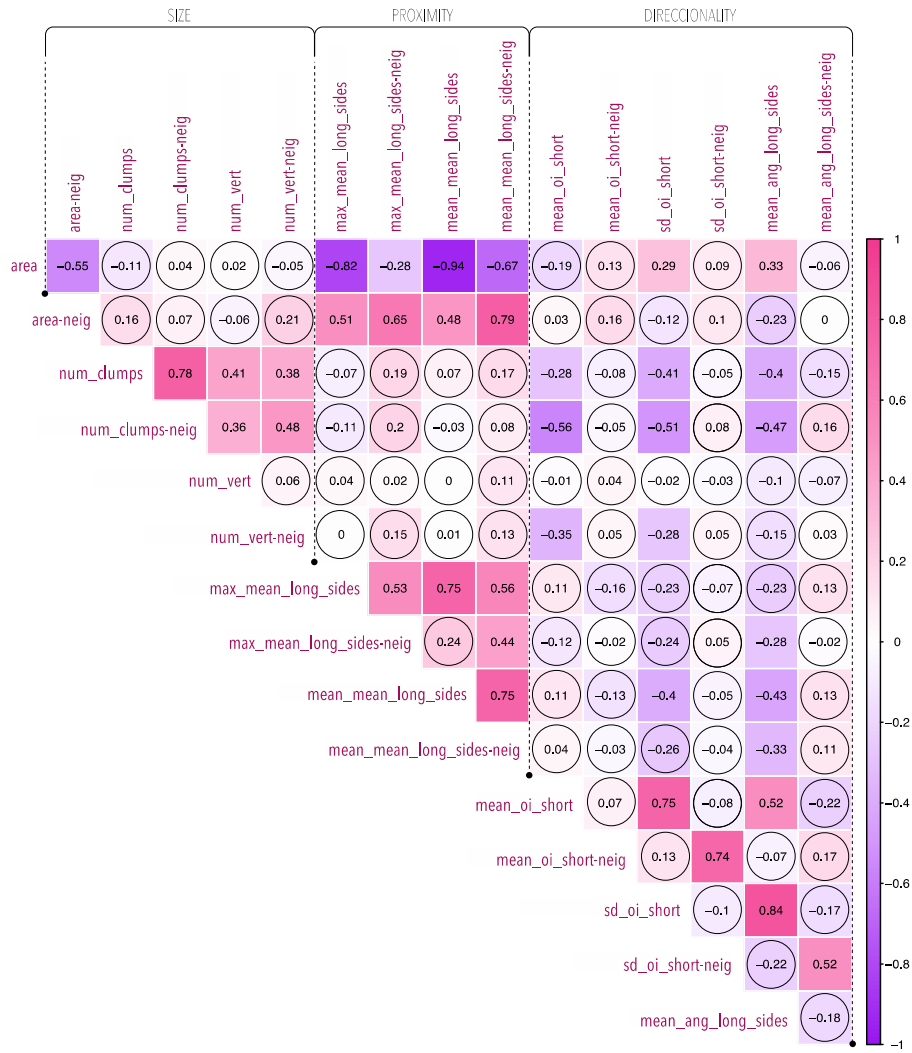

Fig. 15. Sixteen low correlated predicted building metrics from the building level including the neighbouring metrics. Values without circle refers to-significant *p*-values, $p < 0.05$.

Therefore, to proceed with the "reference clustering", eight metrics were selected from the digitised footprint metrics: "area_sum", "inn_irr_angle_max", "oi_mean_mean", "dvl_mean_mean"; and their corresponding difference metrics with the neighbours: "area_sum-neig", "inn_irr_angle_max-neig", "oi_mean_mean-neig", "dvl_mean_mean-neig". From the predicted footprint metrics, we retained eight metrics with significant values along the MMRA: "area", "num_clumps-neig", "max_mean_long_sides-neig", "mean_ang_long_sides", and their corresponding neighbour metrics: "area-neig", "num_clumps", "max_mean_long_sides", "mean_ang_long_sides-neig".

As it can be observed, not all the digitised footprint metrics can be explained by the predicted footprint metrics. The building size metrics ("mean_area", "mean_short", "mean_num_vert") and the inner irregularity metrics ("mean_inn_irr_ang") cannot be captured by metrics in the predicted footprints. This comes from the difficulty of the U-net model to extract individual building footprints consistently.

The optimal number of clusters was determined using the Dunn index. The Dunn index was selected as it showed a robust performance for metrics from both datasets. The optimal number of clusters was three. A hybrid method was chosen, hierarchical-k-means unsupervised clustering, as it produced meaningful clusters for differentiating degrees of deprivation. The three classes of deprivation are colour-coded: class number one (red colour), is named "High deprivation"; class number two (green colour) represents the "Medium deprivation" areas; and class number three (blue colour), refers to "Low deprivation" areas (Figs. 17 and 18).

In Highly deprived areas the Size Indicator reflects the high density per block (labelled in the Fig. 18 as "sum_area"). The density has a high and positive association with the number of buildings and we can assume that according to Act 72 (BCK) the higher density, the smaller the estimated area of each building, and the higher the degree of deprivation. With respect to the Inner Irregularity Indicator, the High deprivation class exhibits the highest values of inner irregularity of the buildings ("max_inn_irr_angle_max"). The BCK does not refer to this metric, but local urban experts and community members agree that building materials influence the shape of the building (i.e., brick or concrete are used in less deprived areas, and their construction techniques require more orthogonality of the walls, so less irregularity in the building internal angles) (Fig. 16). As such, the higher the Inner Irregularity Indicator, the higher the degree of deprivation.

The Proximity Indicator within the Mathare DUA, represented by the "mean_dvl_mean" metric, shows that streets in Mathare are characterised by being narrow (Table 5). Act 18 (BCK) indicates a minimum spacing between buildings of 2.4 m. The High Deprivation class shows the lowest values for proximity between buildings, below the minimum BCK measure (i.e., with an average distance of 2.02 m). Consequently, the smaller the distance between buildings, the larger the degree of deprivation of the area.

Fig. 16. All images within Mathare DUA taken by the author. Left: Buildings with walls made of bricks. Middle: Buildings with walls made of iron-sheets. Top- right: Iron-sheet wall detail. Bottom-right: Brick wall detail.

Table 5
Descriptive statistics of the proximity between buildings (m) measured by "mean_dvl_mean" (i.e., the mean per grid of the mean of the four closest neighbours per building). Each Deprivation class is represented.

Descriptive Statistics	All grids	High Deprivation	Medium Deprivation	Low Deprivation
Min.	0.92	0.92	1.07	6.95
1st Qu.	1.51	1.39	2.88	9.56
Median	2.18	1.80	3.39	12.17
Mean	2.93	2.02	3.79	10.74
3rd Qu	3.37	2.38	4.62	12.64
Max	13.11	4.55	7.87	13.11

The values of the Directionality Indicator ("mean_oi_mean") do not clearly follow the initial hypothesis that the higher the metric, the higher the deprivation, but the opposite. This can be explained by the fact that directionality is only defined by the four neighbouring buildings. In highly deprived areas, due the high density and the proximity, buildings have low chances to settle freely, so the urban pattern is not as organic as parts with greater open spaces. However, it should be noted that this indicator has often been used to differentiate formal areas from deprived areas, as the fabric is much more organic in the latter (Gomes, 2015).

Furthermore, the neighbouring blocks show homogeneity (lower values refer to homogeneity with the block studied) from which we can derive that the Highly deprived areas are surrounded by high degrees of deprivation in the neighbouring blocks (see "sum_area-neig", "max_-inn_irr_angle-neig", "mean_dvl_mean-neig", "mean_oi_mean-neig" in Fig. 18).

The clustering obtained from the predicted footprints was assessed against the "reference clustering", extracted from the digitised footprints. The overall accuracy was 0.71 (proportion of grids correctly classified). The overall F1-score was 0.47, and the F1-score by class was 0.82 for High Deprivation, 0 for Medium Deprivation, and 0.58 for Low Deprivation, which indicates that the model is able to capture high and low deprivation but fails to capture medium deprivation (Fig. 17).

The values metrics (scaled variables) from both datasets were plotted in a box plot graphic to illustrate the performance of each metric through the deprivation cluster classes (Fig. 18).

Top: metrics from the manually delineated set. Bottom: metrics from the U-net output set.

It can be observed, that in the predicted clustering, the metric density ("area"), and the difference in density of the neighbour's ("area_neig") define High deprivation in a very precise way, even better than with the digitised building metric "sum_area". The number of clumps

Fig. 17. Left: 52 grids clustered from the digitised footprint metrics. Right: 52 grids clustered from the predicted footprint metrics.

Fig. 18. Box plot graphs of the morphological metrics of the clustered morphological metrics.

("num_clumps"), compared to the number of digitised footprints (correlated with "area"), show that the model is unable to extract single buildings when they are very close and their roofs overlap. The number of clumps metric ("num_clumps") and the neighbours' metric ("num_clumps_neig") are the metrics characterising the medium deprivation class, which is not well captured by the predicted cluster. Thus, both metrics do not improve the clustering of degrees of deprivation. The distance between buildings ("max_mean_long_sides") also relates to deprivation degrees, along with the difference in distance between neighbours ("max_mean_long_sides-neig"). These metrics do not perform as well as the "mean_dvl_mean" and "mean_dvl_mean-neig" metrics, but in both cases, the classes are well defined. The metric "mean_ang_long_sides" and its difference with the neighbours ("mean_ang_long_sides-neig") does not capture the orientation index variables of the digitised model.

5. Discussion

In LMIC cities, most inhabitants are living in DUAs, but we know very little - or we have very little quantitative information - about them, due to the absence of data (UN-Habitat, 2020; Kuffer et al., 2017). Availability of open spatial data sets, such as OSM, offers opportunities for the study of urban patterns, however, there still remains an insufficient data coverage of DUAs. The development of new tools that bring building footprints up to date is urgently needed.

Manually digitizing has been the traditional building mapping method, but it is time costly and doesn't allow for frequent updates. This research develops a model based on deep learning techniques (U-Net architecture) designed to cover the gap in DUAs. U-Net has proved successful for extracting building footprints, achieving an F1-score of 0.84, and by visual interpretation examination, getting good coverage of the urban pattern. However, individual segmentation of overlapping buildings or close buildings (less than 60 cm) is still challenging, as the building edges are not well defined (i.e., iron sheets roofs exhibit

irregularities as are made up of several sheets) and the materials of all roofs are similar, making it difficult to distribute pixels (30 cm) into the appropriate classes. The model needs to introduce separation borders between overlapping buildings so the smallest buildings and the densest areas can extract buildings individually. Such a semi-automatic building footprint extraction could be also a first step in producing a building map by manually splitting the building clumps if this would be required.

Also, for future research, it would be interesting to compare the selected architecture, U-Net, with other representative deep neural network architectures (Alom et al., 2019; Bianco, Cadene, Celona, & Napoletano, 2018), such as AlexNet, VGG, DenseNet, GoogLeNet, or FractalNet to analyse their performance for DUAs building segmentation.

The resulting outputs show robust results from the transferability to other nine DUAs of the city. A global application (transferability to other cities) has not been tested as more WV-3 images would be required. Further research will aim to provide a trained model by adding DUAs from other cities, as urban patterns differ from cities (Brown, 1997). This will help to create a global tool to analyse the morphology of DUAs. The ability to map building footprints in deprived areas at large scale will bridge one of the urgent data gaps for global models (e.g., population models), but it is also of significant relevance for local applications such as supporting based data for urban renewal.

Furthermore, the design of urban analytical tools based on morphological metrics, being easily reproducible, is needed. We categorise a deprivation classification approach based on morphological analysis, reflecting area-level deprivation. Building materials (McCartney & Krishnamurthy, 2018), building height (3D analysis) (Taubenböck et al., 2018) or other physical features are recommended for further analysis. Understanding the morphology of **urban deprivation** is a necessary step towards characterizing urban poverty. This has not previously been well covered, due to the lack of data. Neither are global morphological welfare standards, and outdated codes, such as the Building Code of Kenya, are not sufficiently developed documents to categorize deprivation. In addition, it would be needed to go further into the predicted metrics (metrics from the clumps) and provide more useful metrics from them.

Modelling deprivation based on the clustering of morphological features is complex and may also be subjective when identifying deprivation patterns. The morphological metrics only capture the physical characteristics and lack other relevant aspects of deprivation (e.g., socio-economic characteristics). Even if some of the datasets used in this study were generated by local community members, the results must be complemented by robust local knowledge based on community engagement. To fully verify the validity of the morphological characteristics that determine deprivation, the analysis could be linked to other socio-economic variables to understand deprivation in a more multidimensional way (Abascal et al., 2021). Therefore, further research will analyse the transferability of morphological deprivation clustering together with socio-economic indicators in the studied areas, as well as in other areas across the globe. Also, the incorporation of some image-derived features, such as textures and image edges, partially linked from the literature to some facets of deprivation, could add robustness to our degrees of deprivation measured from satellite imagery (Engstrom et al., 2019; Wang, Kuffer, & Pfeffer, 2019).

As we claim, open data is needed. Not only to continue developing tools to characterize deprivation, but also for providing municipalities, NGOs, and communities with missing information. Information could empower stakeholders to advocate on behalf of deprived areas. Mapping poverty should be also discussed from a geo-ethics point of view. In this respect, aggregating metrics at the grid level could be preferable than using disaggregated building metrics. Our aim is that this tool can be merged with other platforms that provide deprivation data at a grid level (Chi, Fang, Chatterjee, & Blumenstock, 2021; Stewart & Oke, 2012; WorldPop, 2020). The implementation of a more common grid size (100 by100 meters grid cell) would make it possible to combine results with other global gridded datasets, e.g., WorldPop (WorldPop, 2020), to derive new spatial indicators of deprivation.

6. Conclusion

This research has generated a method for the reliable identification and delineation of urban deprivation patterns by the characterisation of urban morphology using satellite imagery. The above results, in line with other RS studies, such as DUAs location (Wang, Kuffer, Roy, & Pfeffer, 2019), characterisation (Taubenböck & Kraff, 2014; Georganos et al., 2021) and categorization (Taubenböck et al., 2018) contribute to the improvement of existing data gaps of the most deprived urban areas (Taubenböck, Kraff, & Wurm, 2018b). The extent of our research significantly recalls the scope and limitations of the urban morphology extracted from the satellite, and to which extent degrees of urban deprivation can be defined. Deep learning - U-Net architecture - can extract building footprints in DUAs using VHR imagery - VW3 imagery - showing good performance with an F1 score of 0.84 and a Jaccard index of 0.73 in the testing set. Thus, our results show the ability of U-Net to extract building footprints in diverse DUAs of the city of Nairobi. The extraction of individual buildings remains a challenge, due to the complex urban environments, e.g., variety of physiognomies and limited space between buildings, of DUAs. Moreover, it has been proven that morphological characteristics are linked with degrees of deprivation at a grid level. The deep learning model output combined with a morphological analysis has the capability of characterizing urban deprivation. The availability of building footprint data could act as a game-changer. Commonly, city and local-area planning are done without data, and hence planners do not realize the scale of the consequences of infrastructure developments for communities, such as large-scale evictions. Such evictions are very common in Nairobi, as well as in other sub-Saharan African cities. Overlaying detailed building footprints on planned infrastructure development plans would enable developers to understand, at least partially, the scale of the consequences for communities. Moreover, categorizing physical deprivation could allow for the most critical areas to be detected, and for prioritizing upgrading programs accordingly.

Data and software availability

All the collected data, the manually delineated dataset, the deep learning code, and the morphological analysis codes are available from the SLUMAP project website (https://slumap.ulb.be).

Author's contribution

Conceptualisation, A.A., I.R., S.V., R.S., and M.K.; methodology, A.A., I.R., S.V., and M.K.; writing—original draft preparation, A.A., S.V., and M.K.; writing—review and editing, A.A., I.R., S.V., S.G., R.S., E.W., and M.K.; visualization, A.A., and I.R.; project administration, S.V.; funding acquisition, M.K. and S.G. All authors have read and agreed to the published version of the manuscript.

Funding

The research pertaining to these results received financial aid from the Belgian Federal Science Policy (BELSPO) according to the agreement of subsidy no. SR/11/380. (SLUMAP), from the NWO grant number VI. Veni. 194.025 and from the GCRF Digital Innovation for Development in Africa panel (EPSRC Reference: EP/T029900/1). Stefanos Georganos is supported by a Digital Futures postdoctoral fellowship grant.

Declaration of Competing Interest

The authors declare no conflicts of interest.

Data availability

Data will be available in the project webpage.

Acknowledgement

Special thanks to Abhishek Maiti, from ITC, for supporting the deep learning technical part; to Tommaso Cancellario, from UNAV, for supporting solving coding problems; to Nick Mboga, from ULB, for the comments during the reviewing of the manuscript; to Primož Kovačič, CEO at Spatial Collective, for engaging with communities and sharing the data; and to all Community participants from Mathare, especially to Isaac Muasa ("Kaka"), a community leader in Mathare, for sharing ground knowledge and for the involvement in the project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compenvurbsys.2022.101820.

Bibliography

Abascal, A., Rothwell, N., Shonowo, A., Thomson, D. R., Elias, P., Elsey, H., … Kuffer, M. (2021). *"Domains of deprivation framework" for mapping slums, informal settlements, and other deprived areas in LMICs to improve urban planning and policy: A scoping review*. https://doi.org/10.20944/preprints202102.0242.v2

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). https://doi.org/10.3390/rs11111282

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., … Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. *Electronics (Switzerland), 8*(3). https://doi.org/10.3390/electronics8030292

Anurogo, W., Lubis, M. Z., Pamungkas, D. S., Hartono, & Ibrahim, F. M. (2017). A spatial approach to identify slum areas in east Wara sub-districts, South Sulawesi. *IOP Conference Series: Earth and Environmental Science, 98*(1). https://doi.org/10.1088/1755-1315/98/1/012030

APHRC. (2012). *Population and health dynamics in Nairobi's informal settlements: Report of the Nairobi cross-sectional slums survey (NCSS). Nairobi: APHRC, April* (pp. 1–185).

Arimah, B. (2001). Slums as expressions of social exclusion: Explaining the prevalence of slums in African countries. *United Nations Development Programme.*

Ayala, C., Aranda, C., & Galar, M. (2021). Multi-class strategies for joint building footprint and road detection in remote sensing. *Applied Sciences (Switzerland), 11* (18), 1–18. https://doi.org/10.3390/app11188340

Baud, I., Kuffer, M., Pfeffer, K., Sliuzas, R., & Karuppannan, S. (2010). Understanding heterogeneity in metropolitan India: The added value of remote sensing data for analyzing sub-standard residential areas. *International Journal of Applied Earth Observation and Geoinformation, 12*(5), 359–374. https://doi.org/10.1016/j.jag.2010.04.008

Baud, I., Sridharan, N., & Pfeffer, K. (2008). Mapping urban poverty for local governance in an Indian mega-city: The case of Delhi. *Urban Studies, 45*(7), 1385–1412. https://doi.org/10.1177/0042098008090679

Bianco, S., Cadene, R., Celona, L., & Napoletano, P. (2018). Benchmark analysis of representative deep neural network architectures. *IEEE Access, 6*, 64270–64277. https://doi.org/10.1109/ACCESS.2018.2877890

Bishop, C. M. (2006). Pattern recognition and machine learning. In *Springer Science+Business Media*. LLC. Springer.

Brock, G., Susmita, D., Pihur, V., & Datta, S. (2008). clValid: An R package for cluster validation. *Journal of Statistical Software, 25*(4). https://doi.org/10.1016/0038-1098(77)91248-0

Brown, F. E. (1997). Space is the machine. In *, Vol. 18. Design Studies*. https://doi.org/10.1016/s0142-694x(97)89854-7. Issue 3.

Camps-Valls, G., Tuia, D., Xiang-Zhu, X., & Reichstein, M. (2021). *Deep learning for the earth sciences: A comprehensive approach to remote sensing, climate science and geosciences* (Vol. 148).

Chen, K. M., Leu, C. H., & Wang, T. M. (2019). Measurement and determinants of multidimensional poverty: Evidence from Taiwan. *Social Indicators Research, 145*(2), 459–478. https://doi.org/10.1007/s11205-019-02118-8

Chi, G., Fang, H., Chatterjee, S., & Blumenstock, J. E. (2021). Micro-estimates of wealth for all low- and middle-income countries. http://arxiv.org/abs/2104.07761.

Davis, M. (2006). Planet of slums. Paperback edition. London; New York: Verso, 2007. https://search.library.wisc.edu/catalog/9910088803602121.

De Castro, A. (2018). *Slums. Desmontando el concepto. ARQ (Santiago). 98* pp. 80–91). https://doi.org/10.4067/s0717-69962018000100080

DigitalGlobe (2014). Retrieved May 3, 2022, from http://worldview3.digitalglobe.com/.

Duque, J. C., Patino, J. E., Ruiz, L. A., & Pardo-Pascual, J. E. (2015). Measuring intra-urban poverty using land cover and texture metrics derived from remote sensing data. *Landscape and Urban Planning, 135*, 11–21. https://doi.org/10.1016/j.landurbplan.2014.11.009

Engstrom, R., Harrison, R., Mann, M., & Fletcher, A. (2019). Evaluating the relationship between contextual features derived from very high spatial resolution imagery and urban attributes: A case study in Sri Lanka. In *Paper presented at the 2019 Joint Urban Remote Sensing Event (JURSE)*.

Everit, B., & Skrondal, A. (2010). *The Cambridge dictionary of statistics*. Cambridge University Press.

Georganos, S., Abascal, A., Kuffer, M., Wang, J., Owusu, M., Wolff, E., & Vanhuysse, S. (2021). Is it all the same? Mapping and characterizing deprived urban areas using worldview-3 superspectral imagery. a case study in nairobi, kenya. *Remote Sensing, 13*(24). https://doi.org/10.3390/rs13244986

Gilbert, A. (2007). The return of the slum: Does language matter? *International Journal of Urban and Regional Research, 31*(4), 697–713. https://doi.org/10.1111/j.1468-2427.2007.00754.x

Gomes, M. (2015). *The geometry of slums: Boundaries, packing and diversity. July 2001*.

Graesser, J., Cheriyadat, A., Vatsavai, R. R., Chandola, V., Long, J., & Bright, E. (2012). Image based characterization of formal and informal neighborhoods in an urban landscape. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5*(4), 1164–1176. https://doi.org/10.1109/JSTARS.2012.2190383

Grippa, T., Georganos, S., Zarougui, S., Bognounou, P., Diboulo, E., Forget, Y., Lennert, M., Vanhuysse, S., Mboga, N., & Wolff, E. (2018). Mapping urban land use at street block level using OpenStreetMap, remote sensing data, and spatial metrics. *ISPRS International Journal of Geo-Information, 7*(7). https://doi.org/10.3390/ijgi7070246

Henninger, N., & Snel, M. (2002). Where are the poor? Experiences with the development and use of poverty maps. In. *Urban Studies, 45*, Issue 7.

Huchzermer, M. (2011). *Cities with 'slums': From informal settlement eradication to a right to the City in Africa*. UCT Press. https://doi.org/10.1111/juaf.12008

Ibrahim, M. R., Haworth, J., & Cheng, T. (2021). URBAN-i: From URBAN scenes to mapping slums, transport modes, and pedestrians in cities using deep learning and computer vision. *Environment and Planning B: Urban Analytics and City Science, 48*(1), 76–93. https://doi.org/10.1177/2399808319846517

Jochem, W. C., & Tatem, A. J. (2021). Tools for mapping multi-scale settlement patterns of building footprints: An introduction to the R package foot. *PLoS One, 16*(2 February), 1–19. https://doi.org/10.1371/journal.pone.0247535

Joshi, P., Sen, S., & Hobson, J. (2002). Experiences with surveying and mapping Pune and Sangli slums on a geographical information system (GIS). *Environment and Urbanization, 14*(2), 225–240. https://doi.org/10.1177/095624780201400218

K'Akumu, O. A., & Olima, W. H. A. (2007). The dynamics and implications of residential segregation in Nairobi. *Habitat International, 31*(1), 87–99. https://doi.org/10.1016/j.habitatint.2006.04.005

Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* (pp. 1–15).

KNBS. (2019). Distribution of population by administrative units. In *, Vol. II. 2019 Kenya Population and Housing Census*. http://www.knbs.or.ke.

Kohli, D., Sliuzas, R., Kerle, N., & Stein, A. (2012). An ontology of slums for image-based classification. *Computers, Environment and Urban Systems, 36*(2), 154–163. https://doi.org/10.1016/j.compenvurbsys.2011.11.001

Kohli, D., Sliuzas, R., & Stein, A. (2016). Urban slum detection using texture and spatial metrics derived from satellite imagery. *Journal of Spatial Science, 61*(2), 405–426. https://doi.org/10.1080/14498596.2016.1138247

Krishna, A., Sriram, M. S., & Prakash, P. (2014). Slum types and adaptation strategies: identifying policy-relevant differences in Bangalore. *Environment and Urbanization, 26*(2), 568–585. https://doi.org/10.1177/0956247814537958

Krizhevsky, A., Sutskever, I., Hinton, E., & G.. (2015). ImageNet classification with deep convolutional neural networks. *Handbook of Approximation Algorithms and Metaheuristics, 1–1432*. https://doi.org/10.1201/9781420010749

Kuffer, M., & Barros, J. (2011). Urban morphology of unplanned settlements: The use of spatial metrics in VHR remotely sensed images. *Procedia Environmental Sciences, 7*, 152–157. https://doi.org/10.1016/j.proenv.2011.07.027

Kuffer, M., Barros, J., & Sliuzas, R. V. (2014). The development of a morphological unplanned settlement index using very-high-resolution (VHR) imagery. *Computers, Environment and Urban Systems, 48*, 138–152. https://doi.org/10.1016/j.compenvurbsys.2014.07.012

Kuffer, M., Pfeffer, K., & Sliuzas, R. (2016). Slums from space-15 years of slum mapping using remote sensing. *Remote Sensing, 8*(6). https://doi.org/10.3390/rs8060455

Kuffer, M., Pfeffer, K., Sliuzas, R., Baud, I., & van Maarseveen, M. (2017). Capturing the diversity of deprived areas with image-based features: The case of Mumbai. *Remote Sensing, 9*(4). https://doi.org/10.3390/rs9040384

Kuffer, M., Thomson, D. R., Boo, G., Mahabir, R., Grippa, T., Vanhuysse, S., … Kabaria, C. (2020). The role of earth observation in an integrated deprived area mapping "system" for low-to-middle income countries. *Remote Sensing, 12*, 982. https://doi.org/10.3390/rs12060982

Kuffer, M., Wang, J., Thomson, D. R., Georganos, S., Abascal, A., Owusu, M., & Vanhuysse, S. (2021). Spatial information gaps on deprived urban areas (slums) in low-and-middle-income-countries: A user-centered approach. In urban. *Science, 5* (Issue 4). https://doi.org/10.3390/urbansci5040072

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer.

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

Li, L., Wang, C., Zhang, H., & Zhang, B.. *Residual UNet for urban building change detection with sentinel-1 SAR data*. 1498–1501. https://ieeexplore.ieee.org/xpl/conhome/8891871/proceeding.

Li, P., Wang, D., Wang, L., & Lu, H. (2018). Deep visual tracking: Review and experimental comparison. *Pattern Recognition, 76*, 323–338. https://doi.org/10.1016/j.patcog.2017.11.007

Liang, D., De Jong, M., Schraven, D., & Wang, L. (2021). Mapping key features and dimensions of the inclusive city: A systematic bibliometric analysis and literature study. *International Journal of Sustainable Development and World Ecology*. https://doi.org/10.1080/13504509.2021.1911873

Lilford, R., Kyobutungi, C., Ndugwa, R., Sartori, J., Watson, S. I., Sliuzas, R., … Ezeh, A. (2019). Because space matters: Conceptual framework to help distinguish slum from non-slum urban areas. *BMJ Global Health, 4*(2). https://doi.org/10.1136/bmjgh-2018-001267

Liu, R., Kuffer, M., & Persello, C. (2019). The temporal dynamics of slums employing a CNN-based change detection approach. *Remote Sensing, 11*(23). https://doi.org/10.3390/rs11232844

Local Government. (1968). *Building Code* (Vol. 15). https://doi.org/10.1007/978-1-4842-5343-4_6. Issue 15.

Mason, S. O., & Fraser, C. S. (1998). Image sources for informal settlement management. *The Photogrammetric Record, 16*(92), 313–330. https://doi.org/10.1111/0031-868X.00128

McCartney, S., & Krishnamurthy, S. (2018). Neglected? Strengthening the morphological study of informal settlements. *SAGE Open, 8*(1), 1–11. https://doi.org/10.1177/2158244018760375

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika, 50*(2), 159–179. https://doi.org/10.1007/BF02294245

Morgan, W. T. W. (1967). *Nairobi*. City and Region: Oxford University Press.

Mumford, L. (1961). *The City in history: Its origins, its transformations and its prospect*. Penguin Books.

Naik, N., Raskar, R., & Hidalgo, C. A. (2016). Cities are physical too: Using computer vision to measure the quality and impact of urban appearance. *American Economic Review, 106*(5), 128–132. https://doi.org/10.1257/aer.p20161030

Oberay, A. S. (1993). *Population growth, employment and poverty in third world and mega-cities. In explorations in economic history*. St. Martin'S Press. https://doi.org/10.1080/00033799300200371

Pan, Z., Xu, J., Guo, Y., Hu, Y., & Wang, G. (2020). *Deep learning segmentation and classification for urban villages using Worldview satellite image based on U-net* (pp. 1–18).

Peng, D., Zhang, Y., & Guan, H. (2019). End-to-end change detection for high resolution satellite images using improved UNet++. *Remote Sensing, 11*(11). https://doi.org/10.3390/rs11111382

Ranguelova, E., Weel, B., Roy, D., Kuffer, M., Pfeffer, K., & Lees, M. (2018). Image based classification of slums, built-up and non-built-up areas in Kalyan and Bangalore, India. *Eur. J. Remote Sens., 1-22*. https://doi.org/10.1080/22797254.2018.1535838

Rastogi, K., Bodani, P., & Sharma, S. A. (2020). Automatic building footprint extraction from very high-resolution imagery using deep learning techniques. *Geocarto International, 6049*. https://doi.org/10.1080/10106049.2020.1778100

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *, 9351. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_28

Schirmer, P. M., van Eggermond, M. A. B., & Axhausen, K. W. (2014). The role of location in residential location choice models: A review of literature. *Journal of Transport and Land Use, 7*(2), 3–21. https://doi.org/10.5198/jtlu.v7i2.740

SCollective. (2020). GitHub-Datasets/mathare at master- SCollective/Datasets.. https://github.com/SCollective/Datasets/tree/master/mathare.

Simon, H. A. (1962). The architecture of complexity. *The Roots of Logistics, 106*(6), 335–361. https://doi.org/10.1007/978-3-642-27922-5_23

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* (pp. 1–14).

Sirko, W., Kashubin, S., Ritter, M., Annkah, A., Bouchareb, Y. S. E., Dauphin, Y., … Quinn, J. (2021). *Continental-scale building detection from high resolution satellite imagery* (pp. 1–15). http://arxiv.org/abs/2107.12283.

Sliuzas, R., & Kuffer, M. (2008a). Analysing the spatial heterogeneity of poverty using remote sensing: Typology of poverty areas using selected RS based indicators. In *Proceedings of the EARSel Joint Workshop: Remote Sensing New Challenges of High Resolution, i* (pp. 158–167).

Sliuzas, R., Mboup, G., & de Sherbinin, A. (2008). *Report of the expert group meeting on slum identification and mapping* (p. 36). UN-Habitat, ITC, January: Report by CIESIN.

Sliuzas, R. V., & Kuffer, M. (2008b). Analysing the spatial heterogeneity of poverty using remote sensing: Typology of poverty areas using selected RS based indicators. In *Proceedings of the EARSel Joint Workshop: Remote Sensing New Challenges of High Resolution*.

Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society, 93*(12), 1879–1900. https://doi.org/10.1175/BAMS-D-11-00019.1

Sun, W., Chen, B., & Messinger, D. W. (2014). Nearest-neighbor diffusion-based pan-sharpening algorithm for spectral images. *Optical Engineering, 53*(1), Article 013107. https://doi.org/10.1117/1.oe.53.1.013107

Taubenböck, H., & Kraff, N. J. (2014). The physical face of slums: A structural comparison of slums in Mumbai, India, based on remotely sensed data. *Journal of Housing and the Built Environment, 29*(1), 15–38. https://doi.org/10.1007/s10901-013-9333-x

Taubenböck, H., Kraff, N. J., & Wurm, M. (2018). The morphology of the Arrival City - a global categorization based on literature surveys and remotely sensed data. *Applied Geography, 92*(September 2017), 150–167. https://doi.org/10.1016/j.apgeog.2018.02.002

Taubenböck, H., Kraff, N. J., & Wurm, M. (2018b). The blind spot – Reducing knowledge gaps in urban poverty with earth observation interacting with structured and unstructured geodata | science trends. https://sciencetrends.com/the-blind-spot-reducing-knowledge-gaps-in-urban-poverty-with-earth-observation-interacting-with-structured-and-unstructured-geodata/.

Thomson, D., Kuffer, M., Boo, G., Hati, B., Grippa, T., Elsey, H., … Sliuzas, R. (2020). Need for an integrated deprived area "slum" mapping system (IDEAMAPS) in low- and middle-income countries (LMICs). *Social Sciences, 9*(80), 17. https://doi.org/10.3390/socsci9050080

UN-Habitat. (2003). The challenge of slums. In *, Vol. 238. Earthscan Publications Ltd*. https://doi.org/10.1006/abio.1996.0254. Issue 1.

UN-Habitat. (2006). *Nairobi Urban Sector Profile*.

UN-Habitat. (2015). *Sustainable development goals* (Vol. 148).

UN-Habitat. (2020). *Un-Habitat Sub-saharan Africa Atlas*.

UN-Habitat. (2020). World Cites Report 2020. The Value of Sustainable Urbanization. United Nations Human Settlements Programme (UN-Habitat). 10.18356/c41ab67e-en.

UN-Habitat, Statistics-Division, U, & Alliance, C. (2002). Expert Group Meeting on Urban Indicators Secure Tenure, Slums and Global Sample of Cities. November, 33. http://www.citiesalliance.org/sites/citiesalliance.org/files/expert-group-meeting-urban-indicators%5B1%5D.pdf.

Van Zwanenberg, R. (1972). *History and theory of urban poverty in Nairobi. The problem of Slum Development*. Institute for Development Studies, University of Nairobi.

Venerandi, A., Quattrone, G., & Capra, L. (2018). A scalable method to quantify the relationship between urban form and socio-economic indexes. *EPJ Data Science, 7*(1), 1–21. https://doi.org/10.1140/epjds/s13688-018-0132-1

Wang, J., Kuffer, M., & Pfeffer, K. (2019). The role of spatial heterogeneity in detecting urban slums. *Computers, Environment and Urban Systems, 73*, 95–107. https://doi.org/10.1016/j.compenvurbsys.2018.08.007

Wang, J., Kuffer, M., Roy, D., & Pfeffer, K. (2019). Deprivation pockets through the lens of convolutional neural networks. *Remote Sensing of Environment, 234*, Article 111448. https://doi.org/10.1016/j.rse.2019.111448

White, R., Engelen, G., & Uljee, I. (2015). *Modeling cities and regions as complex systems: From theory to planning applications*. The MIT Press. https://doi.org/10.7551/mitpress/9780262029568.001.0001

World Bank. (2008). Approaches to urban slums. In *Approaches to Urban Slums*. https://doi.org/10.1596/978-0-8213-7354-5

World Bank. (n.d.). Population living in slums - sub-Saharan Africa. Retrieved September 29, 2021, from https://data.worldbank.org/indicator/EN.POP.SLUM.UR.ZS?locations=ZG.

WorldPop. (2020). WorldPop: Population Counts. https://www.worldpop.org/project/categories?id=3.

Wu, G., Guo, Z., Shi, X., Chen, Q., Xu, Y., Shibasaki, R., & Shao, X. (2018). A boundary regulated network for accurate roof segmentation and outline extraction. *Remote Sensing, 10*(8), 1–19. https://doi.org/10.3390/rs10081195

Wurm, M., Stark, T., Zhu, X. X., Weigand, M., & Taubenböck, H. (2019). Semantic segmentation of slums in satellite images using transfer learning on fully D. M. S neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing, 150*, 59–69. https://doi.org/10.1016/j.isprsjprs.2019.02.006

Wurm, M., & Taubenböck, H. (2018). Detecting social groups from space – Assessment of remote sensing-based mapped morphological slums using income data. *Remote Sensing Letters, 9*(1), 41–50. https://doi.org/10.1080/2150704X.2017.1384586

Xing, J., Ruixi, Z., Zen, R., Arsa, D. M. S., Khalil, I., & Bressan, S.. Building extraction from google earth images. https://doi.org/10.1145/3366030.3368456.

Yang, S. (2018). *How to extract building footprints from satellite images using deep learning*. Microsoft Azure. https://azure.microsoft.com/es-es/blog/how-to-extract-building-footprints-from-satellite-images-using-deep-learning/.

Zenodo repository with building footprint and U-Net code for building segmentation in the deprived urban areas (slums) of Nairobi, Kenya. https://zenodo.org/record/6514859#.YnFQv9pBzIV. Zenodo.

Zhong, M., Zhang, Y., Beykaei, A., Salehi, B., Ircha, M., Gweon, Y., & Gao, S. (2010). Extracting land cover/use from remotely sensed imagery: Potentials for urban planning. In *Proceedings of the Conference on Traffic and Transportation Studies, ICTTS, 383(March 2015)* (pp. 1–12). https://doi.org/10.1061/41123(383)1