Towards an Indicator-Based Morphological Informality Model for Sub-Saharan Africa Using Open Building Footprint and Road Data (Version 1)

Sebastian Hafner[1], Qunshan Zhao[1], Angela Abascal[2], Manuella Comerio de Paulo[1], Grant Tregonning[1], Alexandra Middleton[3], Adenike Shonowo[1], Monika Kuffer[4], Ryan Engstrom[5] , Dana R. Thomson[6], Francis C. Onyambu[7], Caroline Kabaria[7], Peter Elias [8], Oluwatoyin Odulana[8], Bunmi Alugbin[8], Kehinde Baruwa[8], João Porto de Albuquerque[1]

[1] Urban Big Data Centre (UBDC), University of Glasgow, Glasgow, United Kingdom
[2] Engineer Department, Public University of Navarra, Pamplona, Spain
[3] University of Warwick, Warwick, United Kingdom
[4] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, Netherlands
[5] Department of Geography & Environment, George Washington University, Washington, D.C., United States
[6] Center for Integrated Earth System Information (CIESIN), Columbia University, New York, United States
[7] African Population and Health Research Center (APHRC), Nairobi, Kenya
[8] University of Lagos, Lagos, Nigeria

Keywords: SDG 11; Slums; Unplanned Urbanization; Urban Morphometrics; Urban Deprivation.

Abstract

This study addresses the challenge of accurately mapping informal settlements, which are home to over a billion people globally. Current maps often simplify these areas into binary categories, ignoring the nuanced dimensions of deprivation. The research focuses on ”unplanned urbanization,” a key domain in informal settlement mapping, and proposes a method to classify morphological informality into three deprivation levels (low, medium, and high) based on two subdomains: small, dense structures (SDS) and irregular settlement layouts (ISL). The methodology involves analyzing building footprints and road network data using urban morphometrics, clustering these metrics into subdomains with k-means, and validating results with community-sourced reference data. Tested in Nairobi, Kenya, and Lagos, Nigeria, the model achieves good performance (F1 $>$ 65 for indicator maps) but faces challenges in the medium informality class, particularly in Nairobi, where community feedback diverges significantly. Despite an overall accuracy of 48 % for Nairobi and 60 % for Lagos, the model offers a framework for continuous improvement. This work highlights the value of integrating local perspectives into mapping efforts and provides a scalable, transferable approach for identifying levels of morphological informality.

1. Introduction

Over one billion people globally are estimated to reside in slums, informal settlements, or other deprived areas. In Sub-Saharan Africa, UN-Habitat estimates that approximately half of the population lived in slums in 2020, with this proportion expected to rise. Despite the critical need for accurate mapping of deprived areas, available maps are often outdated or overly simplistic, typically offering a binary distinction between slums and formal areas (Kuffer et al., 2024). However, mapping informal settlements beyond this dichotomy is challenging (Kamalipour, 2016). In particular, while most formal and informal morphologies are relatively easily identified, mixed morphologies are far more problematic to differentiate (Dovey and Kamalipour, 2017).

Contributing to the challenge of deprived area mapping is that various terms aimed at the settlements of the urban poor are used inconsistently in the literature (e.g., slum, informal settlement, and ghettos). This has led to imprecise target classes and varying measurement methods for deprived area mapping (Taubenböck et al., 2018). In particular, due to the complexity and multidimensionality of the concept of deprivation, it is important to use precise definitions for what is being mapped. Abascal et al. (2022b) disentangled deprivation into nine distinctive domains based on a large scoping review (Abascal et al., 2022b). This framework developed by the Integrated Deprived Area Mapping System (IDEAMAPS) Network calls for specialized mapping efforts focusing on individual domains of deprivation, acknowledging that data requirements vary significantly across the domains.

The unplanned urbanization domain of deprivation is associated with rapid and unplanned in-migration to an area. Unplanned urbanization can result in tightly packed and unplanned housing, limited green space, and a lack of roads (Abascal et al., 2022b). Therefore, the domain is also associated with the morphology of informal settlements which are generally characterized by dense, small-grain, and irregular urban morphology (Kamalipour and Dovey, 2019). The unplanned urbanization domain can be measured with volunteered geographic information such as OpenStreetMap (OSM) or Earth observation (EO) data. Indeed, several works utilized open building footprint data, sourced from OSM or extracted from EO data via deep learning techniques, to analyze the morphology of urban areas in Sub-Saharan Africa (Abascal et al., 2022a; Wang et al., 2023; Li et al., 2023). These approaches employ urban morphometrics to numerically describe the form, structure, and arrangement of buildings. Morphometrics clusters characterized by small, dense buildings or compact, organic layouts were associated with informal settlements (Wang et al., 2023; Li et al., 2023).

While these data-driven studies provide valuable insights, other research emphasizes bottom-up approaches informed by locally relevant subdomains of unplanned urbanization (Kuffer et al., 2024). These subdomains have been collaboratively developed with stakeholders from Nigeria (Lagos and Kano) and Kenya (Nairobi), ensuring alignment with local priorities (Kuffer et al., 2024). Moreover, they address themes identified through workshops involving diverse participants, including community members and representatives from local and national governments.

In this study, we propose mapping morphological informality levels (low, medium, and high) based on two subdomains of unplanned urbanization: small, dense structures (SDS) and irregular settlement layouts (ISL). Both subdomains are modeled using urban morphometrics applied to open building footprint data (Fleischmann, 2019; Sirko et al., 2021).

The following are the main contributions of our study:

- We model morphological informality based on a conceptual framework of interpretable indicators for unplanned urbanization that respond to local priority themes.
- The model addresses the limitations of dichotomous informal settlement mapping approaches by introducing three levels of morphological informality: low, medium, and high.
- The morphological informality model is evaluated using community-sourced reference data in Nairobi, Kenya, and Lagos, Nigeria.

2. Related Work

Multiple studies have focused on mapping the settlements of the urban poor based on their physical characteristics and morphologies. Remote sensing is a popular tool for morphology-based slum mapping since satellite imagery provides detailed information on the physical elements of landscapes, including the physical characteristics of slums (Kuffer et al., 2016; Mahabir et al., 2018). In particular, numerous studies used textual features extracted from high and very-high-resolution (VHR) imagery for informal settlement mapping (Duque et al., 2015; Engstrom et al., 2015; Matarira et al., 2022; Owusu et al., 2024). More recent work also combined EO data with a morphometric approach, delivering a systematic numerical characterization of urban form (Wang et al., 2023; Dibble et al., 2019). For the identification of urban patterns, a set of morphological characters was derived from building footprints extracted from VHR satellite imagery via a deep learning model. The characters were summarized at the building level using an unsupervised clustering approach, resulting in 15 morphological clusters. Clusters characterized by small, dense buildings or compact, organic layouts were associated with informal settlements in Nairobi, Kenya (Wang et al., 2023). Li et al. (2023) mapped slums in 95 cities across Sub-Saharan Africa based on morphological and socio-economic indicators, integrated into an unsupervised classification and tree-based clustering framework. Their findings indicate that features such as building coverage, ratio, density, and size correspond to a high prevalence of slums.

Several studies introduced degrees of morphological informality to account for settlement morphologies characterized by a mixture of formal and informal elements (Duque et al., 2015; Dovey and Kamalipour, 2017; Taubenböck et al., 2018; Abascal et al., 2022a; Debray et al., 2023). In the conceptual work of Dovey and Kamalipour (2017), a two-dimensional typology of informal morphologies is introduced. The typology distinguishes three degrees of informality (informal, mixed, and formal) based on architecture and urban design, which refer to the morphology of buildings and the street network, respectively. In both cases, the key criterion for informality is evidence of bottom-up organization, whereas formal morphologies are associated with top-down organization. In the data-driven work by Taubenböck et al. (2018), on the other hand, the spatial structures of 44 arrival cities were measured based on the spatial pattern of settlements (building density, building orientation, and heterogeneity of the pattern) and the morphology of individual buildings (building size and height). These features were used to develop a continuous index for morphological settlement types. The findings show a large morphologic variety of built environments, with slums having similar physical features across the globe. Morphological building features were also linked to a scale of deprivation in Abascal et al. (2022a). Morphological characters related to urban deprivation were derived from building footprints obtained from VHR satellite imagery via a deep learning model. The morphological characters were clustered into three degrees of deprivation in deprived urban areas. Noteworthy, unlike in Wang et al. (2023), the building footprints were not vectorized but morphological characters were directly derived from the raster data. Another study introduced a continuum of Intensity of Plannedness to account for intermediate categories between planned and unplanned areas (Debray et al., 2023). The Urban structure was operationalized by three structural elements: buildings, morphological units, and streets.

3. Study Areas and Grid

The study areas of this research are Nairobi, Kenya, and Lagos, Nigeria (Figure 1). The administrative area of Nairobi spans 695 km^2 with an estimated population of 4.4 million (Kenya National Bureau of Statistics, 2019). Approximately two-thirds of the city's urban population resides in deprived areas as of 2014 (Wamukoya et al., 2020). Nairobi is also home to Kibera, often regarded as the largest informal settlement in Africa. Lagos is the largest metropolitan area and most urbanized state in Nigeria. Its administrative area covering 3,577 km^2 is home to a population of approximately 26 million, making it the second largest city in Africa (The Lagos Resilience Office, 2020).

To define the study area extents for this research, administrative boundaries can be limiting since they are usually determined by historical, political, or administrative decisions and may not necessarily reflect the actual patterns of human settlement or economic activity. For example, the administrative boundary of Nairobi has remained unchanged since 1963 and is considered outdated due to its urban expansion (Abascal et al., 2022a). Likewise, Lagos has expanded significantly beyond its city administrative boundary and is spreading to its neighboring administrative area (Oyalowo, 2022). Instead, we adopt functional urban areas (FUAs) as extents since they are defined by actual urban sprawl and human activities, encompassing the core city and economically or socially integrated surrounding regions (Dijkstra et al., 2019). We apply a uniform buffer (1 km) to FUAs, ensuring peripheral urban zones are included.

Morphological informality is modeled for the FUAs of Nairobi and Lagos at an approximate 100 × 100 m grid. The grid system is based on the Mollweide projection system, an equal-area

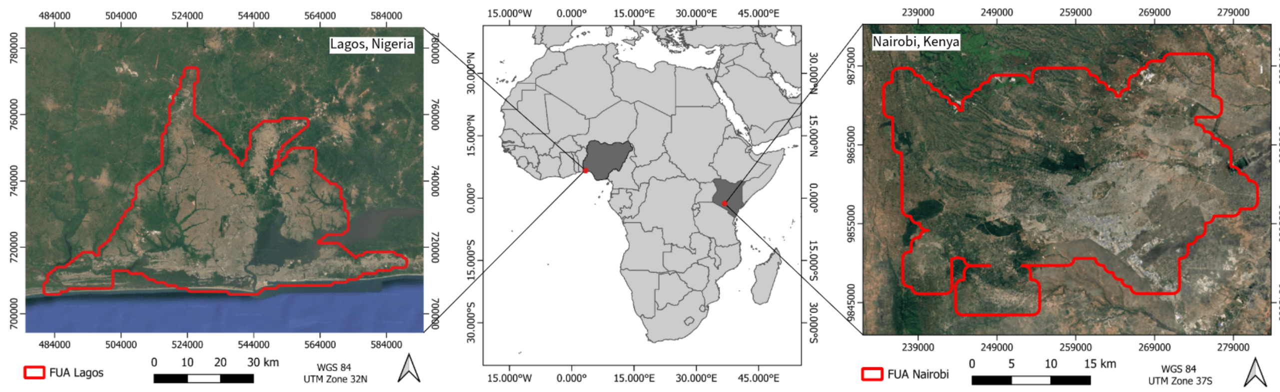


Figure 1. Overview of the study areas.

projection displaying the globe as an ellipse with an axes proportion of 2:1. Due to its suitability for global applications requiring accurate area representations, the grid system is also used by the global human settlement layer (GHSL) data suite. Consequently, this grid facilitates an analysis of gridded population statistics, sourced from the GHSL data suite, within morphological informal areas.

4. Morphological Informality Model

The proposed model identifies three morphological informality levels, low, medium, and high, based on domain indicators (i.e., subdomains of deprivation) for unplanned urbanization. The morphological informality model, including the generation of the indicators, can be summarized in four main steps (Figure 2). Step 1 creates basic urban form elements from open building footprint and road network data. Then, multiple metrics are calculated for the elements to obtain a numerical description of urban form patterns. These element-level metrics are aggregated into a 100 m × 100 m grid using mathematical operators (sum, median, and standard deviation). In Step 3, two sets of aggregated grid-level morphometrics corresponding to the subdomains (SDS and ISL) are selected and clustered using k-means. The resulting morphological clusters are reviewed by local and community experts and categorized as representing either the presence or absence of each subdomain. Finally, Step 4 applies a simple logic to classify cells into three levels of morphological informality: low, medium, and high. The following subsections describe each step in detail.

4.1 Definitions of indicators

Two indicators, SDS (small, dense structures) and ISL (irregular settlement layout), are utilized to map morphological deprivation. These indicators were co-designed with local communities and correspond to subdomains of deprivation that respond to local priority themes and ease communication with stakeholders (Kuffer et al., 2024).

Irregular settlement layout The ISL subdomain focuses on areas characterized by spontaneous, non-linear development, contrasting with more regularly planned areas. Buildings may have unconventional shapes that do not adhere to standard geometric forms often seen in planned developments. Additionally, while buildings are systematically aligned in planned neighborhoods, buildings may be positioned haphazardly with varying distances and orientations to neighboring structures.

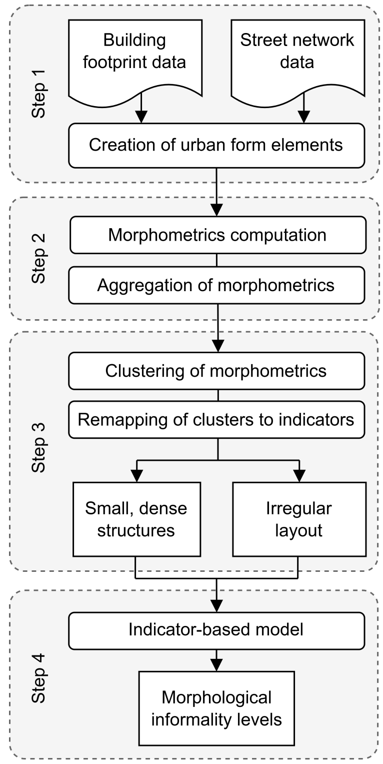


Figure 2. Proposed workflow for the indicator-based morphological informality model.

Small, dense structures The SDS subdomain of deprivation focuses on areas characterized by compact, closely spaced buildings, typically seen in densely populated urban environments. Buildings in these areas may be small and very close to neighboring structures, while buildings in planned neighborhoods are spaced with larger distances.

4.2 Urban form elements

We consider three fundamental urban elements: buildings, morphological tessellation cells, and urban blocks. These elements are often considered to be fundamental categories of morphological analysis (Marshall, 2009).

Buildings The Open Buildings (V3) dataset provided by Google is used as building footprint data (Sirko et al., 2021). This dataset is used as it has greater coverage and completeness within urban areas compared to other openly accessible datasets such as Ecopia, OSM and Microsoft (Chamberlain et al., 2024).

Furthermore, the Open Buildings dataset focuses primarily on classifying buildings within the continent of Africa (and the Global South at large) and is therefore suitable for current and future pilot cities (Sirko et al., 2021).

Tessellation cells Morphological tessellation is a method for consistently partitioning space. Using building footprints and applying the principles of Voronoi tessellation (Fleischmann et al., 2020), morphological tessellations define spatial units called tessellation cells. These are formed considering the influence that each building has on the space surrounding it, through an algorithmic method that captures the spatial configuration of urban form derived from the building footprints (Fleischmann et al., 2022).

Urban blocks Urban blocks are generated based on buildings, their tessellation cells, and the road network. By dissolving the tessellation cells, the blocks serve as an urban form element with a higher-scale aggregation (Fleischmann, 2019). The road networks were retrieved from OpenStreetMap (OSM) (Haklay and Weber, 2008). All accessible pathways and roads available in OSM are included in the road network. It should also be noted that while OSM is a crowdsourced database, its completeness for roads is generally good at a global level (above 80 %), which is also the case for several developing countries (Barrington-Leigh and Millard-Ball, 2017).

4.3 Urban morphometrics

We characterize urban form using a morphological approach by computing a set of urban morphometrics for buildings, tessellation cells, and blocks (Dibble et al., 2019). The selected morphometrics represent urban characters related to the ISL and SDS subdomains of unplanned urbanization. An analysis was performed to determine a comprehensive set of metrics relevant to unplanned urbanization. Specifically, candidates for metrics were first determined based on their ability to capture distinct morphological characteristics of informal settlements using a combination of statistical approaches and a dataset of slum boundaries in Nairobi as references. Thereafter, a multivariate correlation analysis was performed to avoid multiple metrics capturing similar morphological characteristics. Finally, the remaining metrics were rated based on their explainability. The selected metrics, including descriptions and references, are listed in Table 1. A total of 15 parameters were used across the two subdomains, where ISL and SDS are characterized by 7 and 11 parameters, respectively. Three parameters (building adjacency, tessellation cell neighbors, and block count) are shared across the subdomains. For the ISL subdomain, building parameters on orientation, alignment, and adjacency of buildings were selected, as well as the standard deviation of distances between the building centroid and its corners. ISL parameters for the tessellation cells provide information on the neighbors and orientation of the cells. Block count was also used for ISL. On the other hand, for the SDS subdomain, parameters on the area, elongation, neighbor distance, interbuilding distance, and adjacency were used for buildings. For tessellation cells, several area-related parameters (cell area, area ratio, covered area, and equivalent rectangular index) were used, in addition to the number of neighbors.

Since morphometrics are computed at the element level (building, tessellation cell), an aggregation strategy is required to obtain urban morphometrics at the grid level. Figure 3 visualizes the aggregation workflow. Elements are aggregated based on

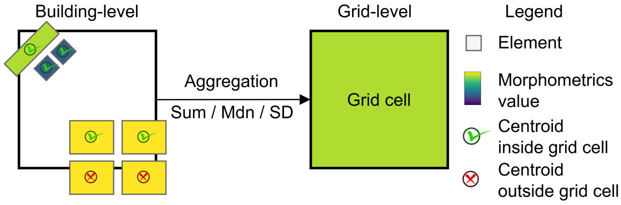


Figure 3. Workflow to aggregate urban morphological metrics calculated at the element level to the grid level. The aggregation is demonstrated for the median operator.

the location of their centroids, where the centroid of the corresponding building is considered for parameters computed for tessellation cells. This is possible due to the one-to-one correspondence between buildings and tessellation cells. It should be noted that since block count is computed at the tessellation level, it can be aggregated via the building correspondence, like the other tessellation-level parameters. Once elements are assigned to a grid cell, different operators, i.e., sum, median, and standard deviation, are used to aggregate parameters (see Table 1). Most parameters are aggregated using the median operator. However, for building area, median and sum are used to get descriptors for not only building size but also the area covered by buildings within the grid cell. The third operator, standard deviation, is exclusively used for the orientation parameter (for buildings and tessellation cells) to describe the regularity of the settlement layout in the ISL subdomain.

4.4 Clustering and remapping of morphometrics

The morphometrics aggregated at the grid level are grouped using a clustering approach to obtain regions with similar urban morphological patterns. K-means with ten clusters was used to cluster the two sets of morphometrics, where the optimal number of clusters was determined using the elbow method. Clusters were then visually interpreted in a co-design activity involving modelers with local knowledge. Each cluster was assigned to either the foreground class (i.e., presence of the subdomain) or background class (i.e., absence of the subdomain). For both subdomains, the foreground class is represented by a single cluster, whereas the other clusters constitute the background.

4.5 Indicator-based typology of informal morphologies

Three morphological informality levels, i.e., low, medium, and high, are derived from the indicators using a simple rule-based model. Figure 4 visualizes the concept of the indicator-based morphological informality model. If the urban morphology is associated with neither of the indicators, the area is modeled as having a low informality level. Notably, a low informality level is also used for areas with no built-up area (i.e., background). On the other hand, areas with one of the indicators present have a medium morphological informality level. Consequently, areas characterized by an irregular settlement layout but with no small, dense structures and areas with small, dense structures but a regular settlement layout are both modeled as medium. Finally, areas characterized by both indicators have a high morphological informality level.

4.6 Model evaluation

The model and indicators are evaluated based on community-sourced reference data, which is available for a limited number of grid cells. This validation data was collected using the

Elem.	Parameter	Description	References	Agg	Indicator ISL	Indicator SDS
Building	Area	Area of building	-	Sum & Mdn	✗	✓
	Elongation	Elongation of the minimum bounding rectangle of a building	Gil et al. (2012)	Mdn	✗	✓
	Centroid corners	Standard deviation of the centroid-corner distances of a building	Schirmer and Axhausen (2016), Cimburova (2017)	Mdn	✓	✗
	Orientation	Orientation deviation of building from cardinal directions in range 0 – 45 degrees	Schirmer and Axhausen (2016)	SD	✓	✗
	Alignment	Median deviation of solar orientation of buildings on adjacent cells from a building	Schirmer and Axhausen (2016)	Mdn	✓	✗
	Neighbor distance	Mean distance to adjacent buildings based on spatial weights	Schirmer and Axhausen (2016)	Mdn	✗	✓
	Interbuilding distance	Mean distance between buildings on adjacent cells based on spatial weights	Fleischmann (2019)	Mdn	✗	✓
	Building adjacency	Ratio of how much buildings tend to join together into larger structures	Fleischmann (2019)	Mdn	✓	✓
Tessellation cell	Area	Area of tessellation cell	-	Mdn	✗	✓
	Equivalent rectangular index	Measure the deviation of a tessellation cell from an equivalent rectangle	Basaraner and Cetinkaya (2017)	Mdn	✗	✓
	Area ratio	Area ratio of building to tessellation cell	Schirmer and Axhausen (2016)	Mdn	✗	✓
	Neighbors	Number of neighbors captured by spatial weights	Hermosilla et al. (2012)	Mdn	✓	✓
	Covered area	Area covered by neighbors defined by spatial weights	-	Mdn	✗	✓
	Orientation	Orientation deviation of tessellation cell from cardinal directions in range 0 – 45 degrees	Schirmer and Axhausen (2016)	SD	✓	✗
Block	Count	Ratio of number of blocks within neighbors defined by spatial weights to the area covered by neighbors	Dibble et al. (2019)	Mdn	✓	✓

Table 1. Urban morphometrics at the grid level comprising the morphological indicators small and dense structures (SDS) and irregular settlement layout (ISL).

Figure 4. The indicator-based typology of informal morphologies used in our model: (bottom-left) Low informality, due to the absence of ISL and SDS, (bottom-right) medium informality, related to the presence of SDS but the absence of ISL, (top-left) medium informality, related to the presence of ISL but the absence of SDS, and (top-right) high, characterized by the presence of ISL and SDS

IDEAMAPS Data Ecosystem platform[1], enabling local communities to validate models for different domains of deprivation. Table 2 lists the number of validated grid cells for the indicators and morphological informality. Grid cells with contradictory validations from multiple participants were removed from the reference data to ensure the reliability and consistency of the ground truth, as conflicting responses could indicate uncertainty or disagreement about the level of deprivation in those locations.

[1] http://www.ideamapsdataecosystem.org/

	Indicator		Morphological Informality		
	ISL	SDS	Low	Medium	High
Nairobi	464	282	56	58	84
Lagos	210	168	50	110	39

Table 2. Number of validation grid cells for the indicators (SDS: small, dense structures and ISL: irregular settlement layout) and the morphological informality levels (low, medium, and high).

Based on this reference data, model outputs are evaluated using the metrics F1 score, intersection over union (IoU), and overall accuracy (Acc), defined as follows:

$$F1 = \frac{TP}{TP + \frac{1}{2} \times (FP + FN)} \tag{1}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{2}$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \tag{3}$$

where TP, TN, FP, and FN denote the number of true positive, true negative, false positive, and false negative pixels, respectively.

https://doi.org/10.5194/isprs-archives-XLVIII-M-7-2025-221-2025

4.7 Implementation details

To calculate the metrics, the momepy package in Python was used (Fleischmann, 2019). This library is part of PySAL (Python Spatial Analysis Library) and is built on top of GeoPandas. OSM data is retrieved using the OSMnx package. The code for the morphological informality model (Version 1) is available on Zenodo[2].

5. Results

We deploy the proposed model in Nairobi, Kenya, and transfer it to Lagos, Nigeria, using Google Open Buildings (v3) data (Sirko et al., 2021). We present model outputs for Nairobi and Lagos in Figures 5a and 5b, respectively. In Nairobi, the vast majority of grid cells were mapped as morphologically formal (96.7 %). Although areas with a medium (1.4 %) or high (1.9 %) level of morphological informality are rare compared to formal areas, the model identified major informal settlements across the city. For example, the model correctly classified Kibera, located towards the southwest of the center of the study site, as an informal settlement. For Lagos, the majority class for grid cells is also low morphological informality (82.5 %). However, a considerable proportion of grid cells are depicted as the medium class (12.9 %), representing a mixture of formal and informal morphologies. Furthermore, more grid cells than in Nairobi are assigned a high morphological informality level (4.6 %). Consequently, large areas of Lagos exhibit non-formal settlement morphologies.

To assess the morphological informality model and the indicators, we used reference data annotated by local community members. The quantitative results are listed in Table 3, and confusion matrices for the informality levels are shown in Figure 6. Indicator maps (ISL and SDS) achieve good performance (F1 > 65) in both cities when compared against the reference data. The morphological informality model (low, medium, and high) also achieves good performance (F1 > 53) for all levels in both cities, except for the medium informality level in Nairobi (F1 $=$ 6). The confusion matrices reveal that most grid cells identified by local communities as medium were incorrectly classified as low by the model (42 grid cells out of the 58 reference cells for medium morphological informality). In Nairobi, many grid cells identified as high informality were also incorrectly classified as low informality (39 grid cells out of 84 reference cells for high morphological informality). In terms of overall accuracy, the models achieve values of 48 % and 60 % for Nairobi and Lagos, respectively.

	Nairobi			Lagos		
	F1 ↑	IoU ↑	Acc ↑	F1 ↑	IoU ↑	Acc ↑
ISL	66.0	49.2	63.6	78.8	65.1	69.0
SDS	79.8	66.4	74.5	74.9	59.8	74.4
Low	55.0	38.0	57.1	53.5	36.6	70.4
Medium	6.3	3.3	70.2	66.0	49.3	65.8
High	58.3	41.2	69.7	56.3	39.2	84.4

Table 3. Quantitative results for indicators (SDS: small, dense structures and ISL: irregular settlement layout) and the morphological informality levels (low, medium, and high) derived from the indicators.

[2] `https://doi.org/10.5281/zenodo.15120032`

6. Discussion

The qualitative and quantitative modeling results demonstrate that the proposed indicator-based morphological informality mapping method is capable of capturing spatial patterns of informality related to unplanned urbanization (Figure 5 and Table 3). Specifically, three levels of morphological informality were distinguished based on the indicators ISL (irregular settlement layout) and SDS (small, dense structures). Our findings are in with several studies reporting that informal settlements can be generally characterized by small, dense buildings or compact, organic layouts (Abascal et al., 2022a; Wang et al., 2023; Li et al., 2023).

6.1 Limitations

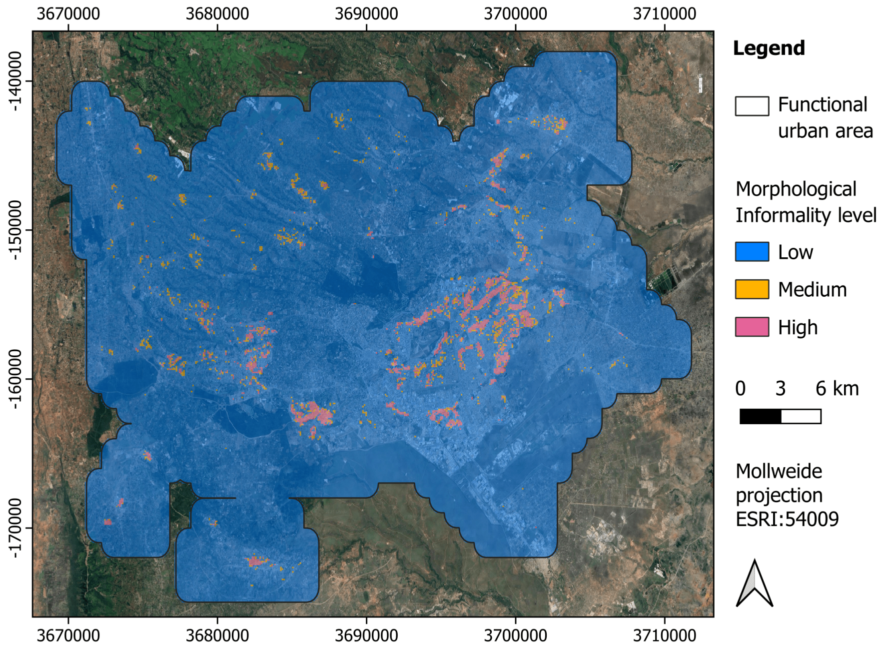
While the model generally achieved promising results across morphological informality levels and indicators, poor performance was obtained for the medium informality level in Nairobi due to confusion with low informality (Figure 6a), which can also be observed in Lagos (Figure 6b). These results indicate that our model underestimates morphological informality levels. Another limitation encountered in this study is using the road network merely as a count of the blocks. Specifically, for the ISL indicator, incorporating analyses that capture the irregularities within the road network could yield more comprehensive insights. The results generally show the challenge of mapping mixed morphologies that are neither completely formal nor informal. These findings are in line with other studies highlighting the challenge of mapping informal settlements beyond the formal-slum dichotomy (Kamalipour, 2016; Dovey and Kamalipour, 2017).

6.2 Perspectives

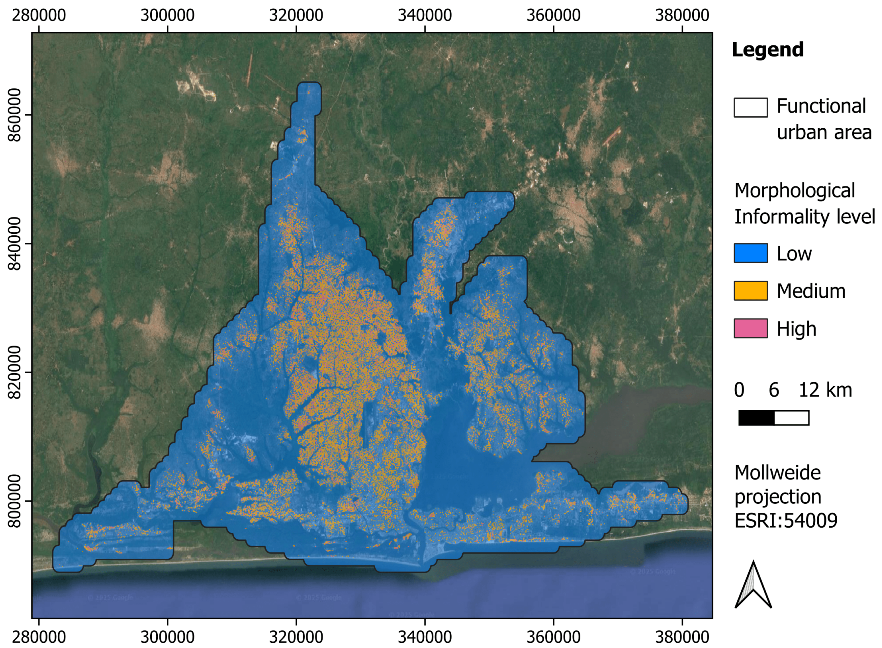
Since the proposed morphological informality model uses open building and road network data, the data quality of these underlying products can have an impact on the model results. With the quality of open data continuously improving, future work will rerun this model with data from Overture Maps, a collaborative open-data initiative launched in 2022. Furthermore, since geographical coverage of informal settlement mapping studies is scarce across the world (Matarira et al., 2022), the model will also be deployed to new cities in Sub-Saharan Africa, Southeast and South Asia, Latin America, and other Global South countries. Finally, the reference data collection efforts using the IDEAMAPS Data Ecosystem Platform will continue and also be scaled up. Consequently, future research should also be able to use these data as labels for deep learning model training since deep learning is emerging as an important technology for slum mapping from EO data.

7. Conclusion

This study presents a first step towards a morphological informality model adhering to conceptually designed subdomains of unplanned urbanization. Furthermore, the model acknowledges the continuous concept of morphological informality by introducing three levels of informality. Our results demonstrate the effectiveness of the proposed model in Nairobi, Kenya, and Lagos, Nigeria, based on community-sourced reference data. However, our results also highlight the challenge of mapping urban areas with mixed morphologies, characterized by both informal and formal elements. Future work will refine the modeling of medium informality, particularly in Nairobi, to better align with local community perspectives.



(a)



(b)

Figure 5. Modeled morphological informality levels for (a) Nairobi, Kenya, and (b) Lagos, Nigeria.

8. Acknowledgment

This work was supported by the Bill & Melinda Gates Foundation INV-045252. The authors also acknowledge the contributions from the stakeholders and partner communities in Lagos (Nigeria), Kano (Nigeria), and Nairobi (Kenya).

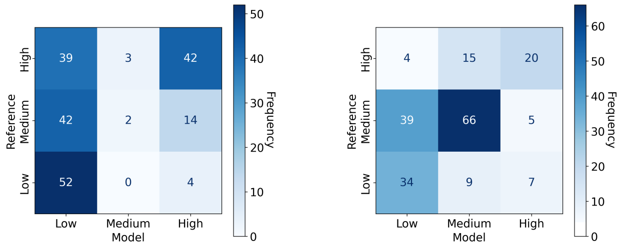
References

Abascal, A., Rodríguez-Carreño, I., Vanhuysse, S., Georganos, S., Sliuzas, R., Wolff, E., Kuffer, M., 2022a. Identifying degrees of deprivation from space using deep learning and morphological spatial analysis of deprived urban areas. *Computers, environment and urban systems*, 95, 101820.

Abascal, A., Rothwell, N., Shonowo, A., Thomson, D. R., Elias, P., Elsey, H., Yeboah, G., Kuffer, M., 2022b. "Domains of deprivation framework" for mapping slums, informal settlements, and other deprived areas in LMICs to improve urban planning and policy: A scoping review. *Computers, environment and urban systems*, 93, 101770.

Barrington-Leigh, C., Millard-Ball, A., 2017. The world's user-generated road map is more than 80% complete. *PloS one*, 12(8), e0180698.

Basaraner, M., Cetinkaya, S., 2017. Performance of shape indices and classification schemes for characterising perceptual shape complexity of building footprints in GIS. *International Journal of Geographical Information Science*, 31(10), 1952–1977.



(a) (b)

Figure 6. Confusion matrices comparing the morphological informality model with community-sourced reference data for (a) Nairobi, Kenya and (b) Lagos, Nigeria.

Chamberlain, H. R., Darin, E., Adewole, W. A., Jochem, W. C., Lazar, A. N., Tatem, A. J., 2024. Building footprint data for countries in Africa: to what extent are existing data products comparable? *Computers, Environment and Urban Systems*, 110, 102104.

Cimburova, Z., 2017. Urban morphology in prague: Automatic classification in GIS. Master's thesis, Czech Technical University.

Debray, H., Kraff, N. J., Zhu, X. X., Taubenböck, H., 2023. Planned, unplanned, or in-between? A concept of the intensity of plannedness and its empirical relation to the built urban landscape across the globe. *Landscape and Urban Planning*, 233, 104711.

Dibble, J., Prelorendjos, A., Romice, O., Zanella, M., Strano, E., Pagel, M., Porta, S., 2019. On the origin of spaces: Morphometric foundations of urban form evolution. *Environment and Planning B: Urban Analytics and City Science*, 46(4), 707–730.

Dijkstra, L., Poelman, H., Veneri, P., 2019. The EU-OECD definition of a functional urban area.

Dovey, K., Kamalipour, H., 2017. Informal/formal morphologies. *Mapping urbanities*, Routledge, 223–248.

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.

Engstrom, R., Sandborn, A., Yu, Q., Burgdorfer, J., Stow, D., Weeks, J., Graesser, J., 2015. Mapping slums using spatial features in accra, ghana. *2015 Joint Urban Remote Sensing Event (JURSE)*, IEEE, 1–4.

Fleischmann, M., 2019. Momepy: Urban morphology measuring toolkit. *Journal of Open Source Software*, 4(43), 1807.

Fleischmann, M., Feliciotti, A., Romice, O., Porta, S., 2020. Morphological tessellation as a way of partitioning space: Improving consistency in urban morphology at the plot scale. *Computers, Environment and Urban Systems*, 80, 101441.

Fleischmann, M., Romice, O., Porta, S., 2022. Applicability of morphological tessellation and its topological derivatives in the quantitative analysis of urban form. *XXVI international seminar on urban form cities as assemblages*.

Gil, J., Beirão, J. N., Montenegro, N., Duarte, J. P., 2012. On the discovery of urban typologies: data mining the many dimensions of urban form. *Urban morphology*, 16(1), 27.

Haklay, M., Weber, P., 2008. Openstreetmap: User-generated street maps. *IEEE Pervasive computing*, 7(4), 12–18.

Hermosilla, T., Ruiz, L., Recio, J., Cambra-López, M., 2012. Assessing contextual descriptive features for plot-based classification of urban areas. *Landscape and Urban Planning*, 106(1), 124–137.

Kamalipour, H., 2016. Forms of informality and adaptations in informal settlements. *International Journal of Architectural Research (ArchNet-IJAR)*, 10(3), 60–75.

Kamalipour, H., Dovey, K., 2019. Mapping the visibility of informal settlements. *Habitat International*, 85, 63–75.

Kenya National Bureau of Statistics, 2019. 2019 kenya population and housing census volume 1: Population by county and sub-county. Technical report, Kenya National Bureau of Statistics.

Kuffer, M., Abascal, A., Engstrom, R., Thomson, D. R., Tregonning, G., Shonowo, A., Zhao, Q., de Albuquerque, J. P., Elias, P., Onyambu, F. C. et al., 2024. Ideamaps: Modelling sub-domains of deprivation with eo and ai. *IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 1562–1566.

Kuffer, M., Pfeffer, K., Sliuzas, R., 2016. Slums from space—15 years of slum mapping using remote sensing. *Remote Sensing*, 8(6), 455.

Li, C., Yu, L., Oloo, F., Chimimba, E. G., Kambombe, O., Asamoah, M., Opoku, P. D., Ogweno, V. W., Fawcett, D., Hong, J. et al., 2023. Slum and urban deprivation in compacted and peri-urban neighborhoods in sub-Saharan Africa. *Sustainable Cities and Society*, 99, 104863.

Mahabir, R., Croitoru, A., Crooks, A. T., Agouris, P., Stefanidis, A., 2018. A critical review of high and very high-resolution remote sensing approaches for detecting and mapping slums: Trends, challenges and emerging opportunities. *Urban Science*, 2(1), 8.

Marshall, S., 2009. Cities design and evolution. *(No Title)*.

Matarira, D., Mutanga, O., Naidu, M., 2022. Texture analysis approaches in modelling informal settlements: A review. *Geocarto International*, 37(26), 13451–13478.

Owusu, M., Nair, A., Jafari, A., Thomson, D., Kuffer, M., Engstrom, R., 2024. Towards a scalable and transferable approach to map deprived areas using Sentinel-2 images and machine learning. *Computers, Environment and Urban Systems*, 109, 102075.

Oyalowo, B., 2022. Implications of urban expansion: land, planning and housing in Lagos. *Buildings & Cities*, 3(1).

Schirmer, P. M., Axhausen, K. W., 2016. A multiscale classification of urban morphology. *Journal of Transport and Land Use*, 9(1), 101–130.

Sirko, W., Kashubin, S., Ritter, M., Annkah, A., Bouchareb, Y. S. E., Dauphin, Y., Keysers, D., Neumann, M., Cisse, M., Quinn, J., 2021. Continental-scale building detection from high resolution satellite imagery. *arXiv preprint arXiv:2107.12283*.

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, 150–167.

The Lagos Resilience Office, 2020. Lagos resilience strategy. Technical report, The Lagos State Government.

Wamukoya, M., Kadengye, D. T., Iddi, S., Chikozho, C. et al., 2020. The Nairobi urban health and demographic surveillance of slum dwellers, 2002–2019: value, processes, and challenges. *Global Epidemiology*, 2, 100024.

Wang, J., Fleischmann, M., Venerandi, A., Romice, O., Kuffer, M., Porta, S., 2023. EO+ Morphometrics: Understanding cities through urban morphology at large scale. *Landscape and Urban Planning*, 233, 104691.