Urban Morphology and Infrastructure Patterns: A LiDAR-Based 3D Cluster Analysis Using Verticalization as a Proxy

Ezequiel Silva Rocha[1], Aluizio Brito Maia[2], Cláudia Maria de Almeida[1], Paulo Roberto da Silva Ruiz[3]

[1]National Institute for Space Research, INPE, Brazil – ezequiel.rocha; claudia.almeida@inpe.br
[2]Aeronautics Institute of Technology, ITA, Brazil – aluizio@ita.br
[3]São Paulo State Technological Colleges, FATEC, Brazil – paulo.ruiz2@fatec.sp.gov.br

Keywords: Verticalization, LiDAR, Self Organizing Maps, Spatial inequality, Belo Horizonte.

Abstract

This study analyzes urban verticalization and socio-spatial patterns in Belo Horizonte using LiDAR-derived building heights and infrastructure indicators. A normalized Digital Surface Model (nDSM) was generated from LiDAR data to map vertical structures with high spatial accuracy. Self-Organizing Maps (SOM) were applied to cluster neighborhoods based on height, infrastructure, and demographic data. The results reveal distinct urban typologies, including central consolidated zones, peripheral vulnerable areas, and hybrid transitional regions. While verticalization reflects urban consolidation, it must be interpreted alongside socio- infrastructural conditions to fully understand spatial inequalities and urban dynamics.

1. Introduction

Urban growth is a dynamic and heterogeneous phenomenon, with cities expanding according to distinct morphological and socio-economic patterns (Mahtta et al., 2019). The complexity of factors driving urban development, ranging from economic opportunities to historical and geographical constraints, results in uneven urban sprawl (Li and Li, 2019). In many cases, this growth is intrinsically linked to social vulnerability and inequality, with certain areas exhibiting lower population density yet offering higher standards of living (Li and Li, 2019).

One of the most prominent indicators of urban transformation is vertical- ization, (Zambon et al., 2019) characterized by the construction of high-rise buildings as a response to limited horizontal space. This process typically occurs in economically valued areas undergoing rapid development and modernization. As such, verticalization often correlates with land value and income levels (Wang et al., 2022), providing a useful proxy for spatial socio- economic classification

Given its relevance, understanding patterns of verticalization is essential for urban planning, infrastructure development, and social research. However, traditional methods based on satellite imagery face limitations due to the absence of height information, restricting the ability to accurately assess building heights and urban density (Shi et al., 2024). The challenge, therefore, lies in integrating vertical data with infrastructure indicators at a detailed scale (Biljecki et al., 2015). To overcome these limitations, Light Detection and Ranging (LiDAR) technology has emerged as a powerful tool, enabling a more precise and comprehensive analysis of urban form through detailed three-dimensional information (Ma et al., 2023; Lao et al., 2021; Bonczak and Kontokosta., 2019).

This study aims to apply an unsupervised machine learning algorithm (Self-Organizing Maps) to quantitatively investigate how urban verticalization, measured through a LiDAR-derived nDSM, relates to socio-spatial patterns of infrastructure and demography in Belo Horizonte and leverage this relationship to produce a classification of urban typologies. By clustering neighborhoods based on this integrated dataset, we seek to reveal a multidimensional perspective on spatial inequality that is not apparent when analyzing these variables in isolation.

2. Materials and Methods

The methodology applied in this study follows the steps out lined in the methodological framework shown in Figure 1.

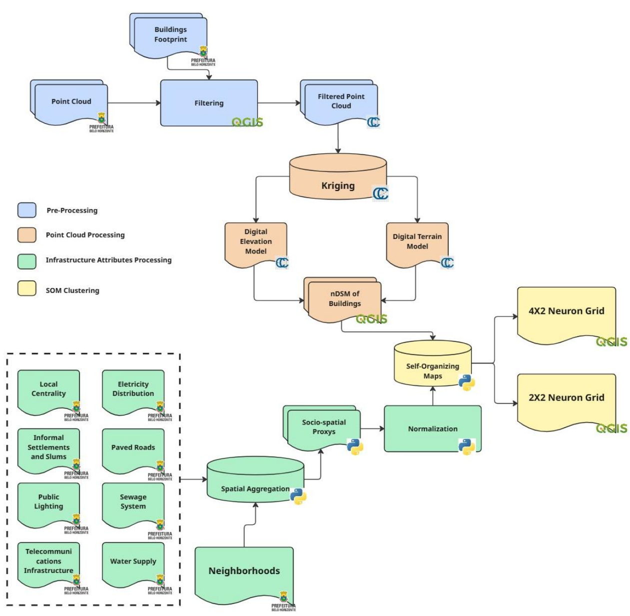

Figure 1. Methodological Procedures.

2.1 Study Area

The study area is located in the state of Minas Gerais, south-eastern Brazil. and corresponds to Belo Horizonte, the state capital, a major urban and economic hub, driving regional development due to its strategic importance. With a population of approximately 2.3 million in the city itself, and over 5 million in the metropolitan region, it stands as the most populous city in Minas Gerais and one of the most influential urban centers in the country.

Belo Horizonte is recognized for its high quality of life and well-developed public infrastructure, ranking among the best capital cities in Brazil. It holds a Human Development Index (HDI) of 0.810 (Prefeitura de Belo Horizonte, 2025a), reflecting strong performance in key social indicators. However, as in many large urban centers, the city faces pronounced socio-spatial inequality (Mahtta et al., 2019). This contrast reinforces the importance of studying verticalization as a way to understand patterns of urban expansion and disparities in infrastructure distribution.

The city was selected for this study due to the availability of comprehensive spatial data. Notably, a complete LiDAR dataset is available for the entire urban area, acquired in 2015. This dataset provides a valuable opportunity to assess urban morphology and vertical growth. Figure 2 shows the location of Belo Horizonte within the state of Minas Gerais in Brazil.

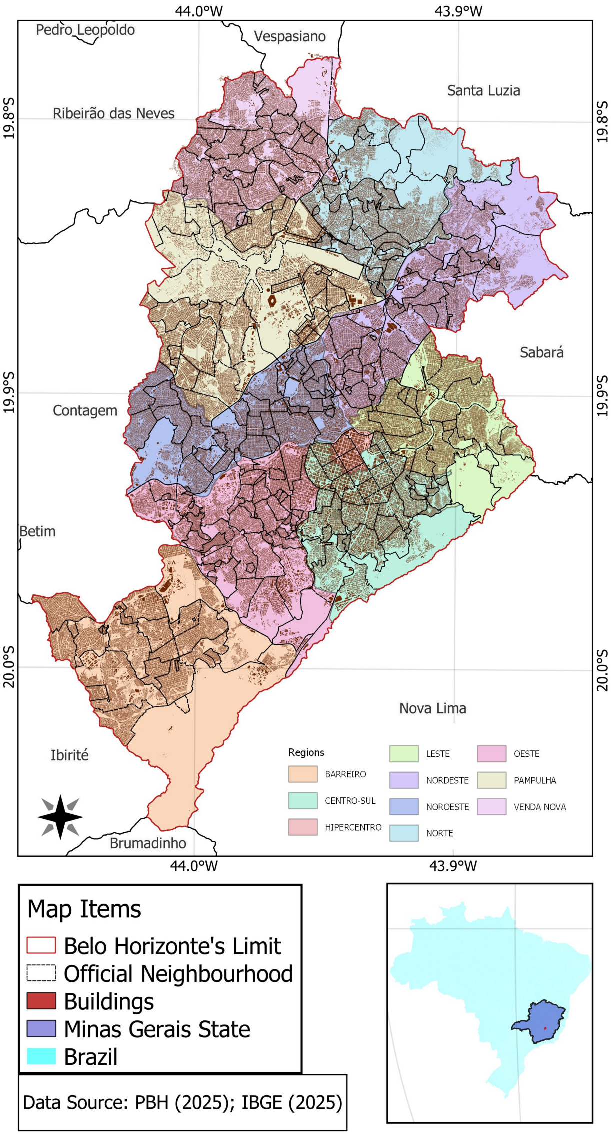

Figure 2. Study Area.

2.2 Data Acquisition and Point Cloud Processing

The data used in this study were obtained from the official open data repository of the Belo Horizonte municipality, which provides both infrastructure datasets and airborne LiDAR point clouds. The LiDAR data were acquired in 2015 using a Leica ALS50-II sensor mounted on an Embraer EMB-810 aircraft, flying at an altitude of 1,750 meters. This configuration yielded a point cloud with an average density of 3 points per square meter (Prefeitura de Belo Horizonte, 2025b).

The raw LiDAR data, originally stored in .xyz format and divided into 397 tiles, was processed using CloudCompare, an open-source software widely adopted in academic and applied remote sensing workflows due to its flexibility and efficient point cloud handling capabilities (Ding et al., 2018; Rajendra et al., 2014; Dewez et al., 2016). The raw point cloud was 55.5 GB in size and contained about 1.68 billion points. An example of a tile is depicted in Figure 3

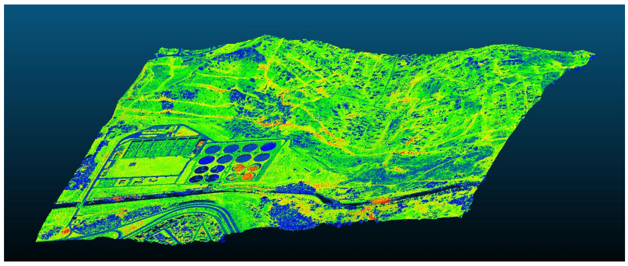

Figure 3. CloudCompare tile.

A Digital Terrain Model (DTM) and a Digital Surface Model (DSM) were generated using Kriging interpolation on CloudCompare, selected for its ability to estimate smooth continuous surfaces from irregularly spaced data points (Goovaerts, 2019). However, the choice of Kriging was made based on its performance in prior urban morphology studies, despite its computational cost and potential smoothing of sharp vertical structures.

To isolate the heights of buildings from other surface features, the nDSM was masked using a building footprint shapefile provided by the Belo Horizonte municipal government (PBH). This procedure, carried out on QGIS, clipped the raster dataset, retaining cell values only within the polygons that correspond to officially demarcated constructed areas.

The normalized Digital Surface Model (nDSM) was computed as the difference between the DSM and the DTM, resulting in a raster product with a spatial resolution of 1 meter. This nDSM was used as a proxy for building height and verticalization patterns throughout the urban landscape.

The main processing steps, including the DTM and DSM generation, took approximately 4 hours and 20 minutes to complete on a workstation with the following configuration: An Intel Core i7-8565U CPU @ 1.99 GHz, 20 GB of RAM, and an NVIDIA GeForce MX110 GPU, running on a Windows 11 operating system.

2.3 Complementary Infrastructure Attributes

To explore the relationship between infrastructure availability and socio-spatial conditions, a set of complementary attributes was computed for each neighborhood, as illustrated in Figure 4. These variables include road density, population living in informal settlements (slums), and indicators of local centrality. All attributes were derived using publicly available data from the open data portal of Belo Horizonte's municipal government (Prefeitura de Belo Horizonte, 2025c).

https://doi.org/10.5194/isprs-annals-X-3-W4-2025-313-2026

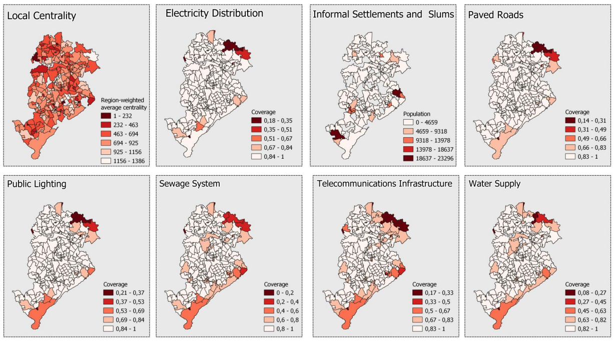

Figure 4. Utilized Attributes.

The processing and spatial aggregation were conducted using Python-based scripts. Infrastructure elements such as streets and public facilities were intersected with the official neighborhood boundaries to calculate their spatial distribution. Population data from slum areas was proportionally allocated to neighborhoods based on area-weighted overlays, accounting for spatial mismatch between informal settlement polygons and neighborhood limits. Similarly, centrality metrics were estimated by calculating weighted averages of infrastructural features intersecting each neighborhood, based on their geometric properties (e.g., length or area). This procedure allowed for the creation of a standardized dataset of socio-spatial indicators suitable for subsequent cluster analysis using Self-Organizing Maps (SOM).

2.4 Self-Organizing Maps

Self-Organizing Maps (SOM) are a type of unsupervised machine learning algorithm designed to project high-dimensional data into a low-dimensional space, while preserving topological relationships within the dataset. Originally introduced by Kohonen (Kohonen, 1982), SOMs use a form of competitive learning to organize input data into clusters represented by neurons arranged in a grid-like structure. In this study, two SOM configurations were tested to explore urban segmentation patterns.

Prior to model training, all variables were normalized using the min–max method, and missing values were imputed using the mean of each respective variable. For the building height attribute, we used the original values as the main explanatory variable, complemented by the socioeconomic covariates. The first configuration used a 2x2 neuron grid, resulting in four coarse-grained clusters, intended to identify broad spatial patterns across the city.
The second configuration employed a finer 4x2 neuron grid, initially producing eight clusters. After examining the resulting U-Matrix visualization (Figure 5), one of the clusters was found to be redundant due to its high similarity with an adjacent group and limited spatial distinction. It was therefore manually merged to the neighboring group, leading to a final configuration of seven clusters.

The U-Matrix is a common visualization technique for Self-Organizing Maps, where each cell represents a neuron and the color scale reflects the average distance between a given neuron and its immediate neighbors. Higher distances, typically shown in darker colors, indicate stronger separability between clusters, while lower distances suggest similarity or continuity in the input data. This technique allows for the visual identification of cluster boundaries and topological transitions within the SOM grid.

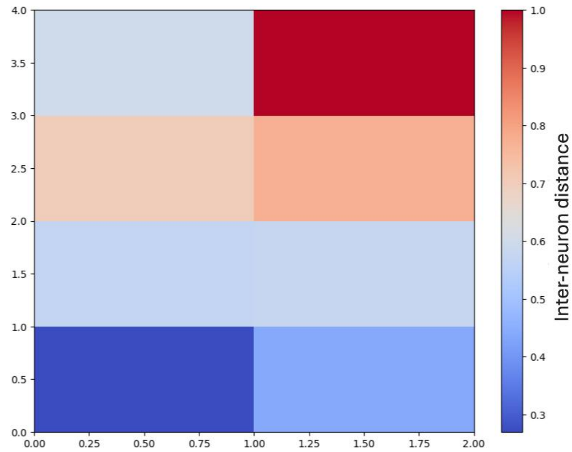

Figure 5. U-Matrix.

This dual-scale clustering approach allowed for a comparative analysis between general and detailed spatial typologies, enabling more nuanced and strategic interpretations of urban dynamics and socio-spatial disparities.

3. Results

3.1 Verticalization Distribution

The spatial distribution of verticalization in Belo Horizonte reflects is shown in Figure 6.

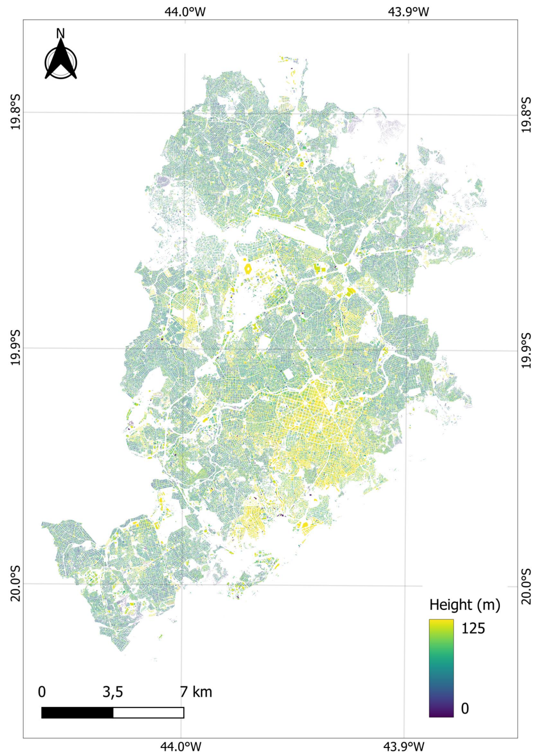

Figure 6. Building Heigh Distribution.

Verticalization of Belo Horizonte exhibits the highest concentration of vertical structures, with building heights ranging from 6 m to 125 m. The tallest structure, the Acácia Building (125 m), was accurately identified by the LiDAR-derived nDSM, confirming the precision and reliability of the dataset used. Most buildings in the central area fall within the range of 20-80 m, showing a consistent pattern of high-rise development within the urban core. Beyond the center, verticalization gradually declines towards the periphery, suggesting a typical radial pattern of urban density, which is a feature commonly observed in Latin American cities (Inostroza, 2017).

A few outliers can be observed in peripheral neighborhoods, likely associated with isolated high-rise developments or topographic distortions due to steep terrain. It is also noteworthy that negative or near-zero values resulting from interpolation artifacts were corrected by truncating them to zero, especially in areas where terrain elevation changes sharply over short distances.

Although verticalization can serve as a proxy for land value or urban consolidation, height alone does not provide a complete picture of socio-spatial dynamics (Harris, 2015). In the following section, we present a clustering analysis that integrates verticalization with infrastructure and demographic indicators to produce a more comprehensive typology of the urban landscape.

3.2 SOM Models Comparation

The 2x2 model was applied and the results are presented in Figure 7.

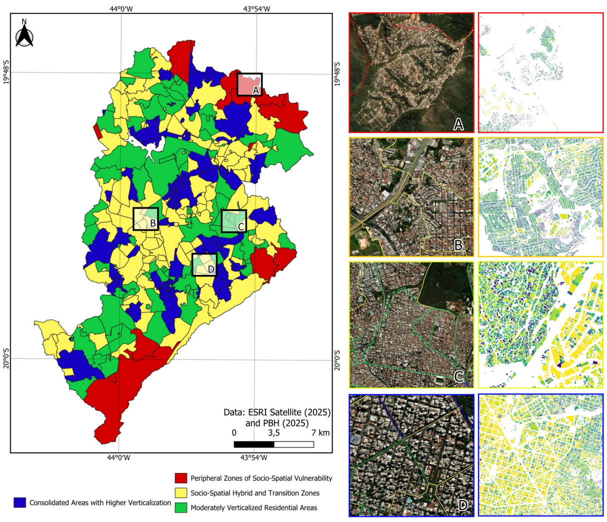

Figure 7. Cluster 2X2.

The results shown in Figure 7A shows the Vitória community, a slum that represents one of the most vulnerable regions of the city, clusterized as peripheral zone of socio-spatial vulnerability, and when it comes to verticalization, it is notable that the region is one of the least developed in this aspect. The region also lacks infrastructural resources and local centralities, embodying the urban traits of the red cluster: Peripheral Vulnerability. In this case, some variables are mutually reinforcing, fewer local centralities are associated with a lack of services attraction, which consequently perpetuates precariousness.

Thus, the lack of verticalization in outskirt areas can be understood as a reflection of informality and exclusion from the formal real estate market. Figure 7B shows the region represented by the Caiçara neighborhood. It is a middle-class neighbor- hood, clusterized as part of the Socio-spatial Hybrid Zones and Transition Areas. Hybrid zones exhibit the coexistence of contrasting elements in the urban landscape and in socio-urban- economic conditions. These areas can include consolidated infrastructure adjacent to informal settlements, or mixed-use typologies combining industrial, commercial, and residential areas.

They can also function as transitional regions, subject to gentrification, urban requalification, or real estate development pressure. Figure 7C illustrates the Sagrada Família neighborhood, a high middle-class area characterized by low levels of verticalization, but notable for its strong local centrality and dense population. Despite being classified under the cluster of moderately verticalized residential areas, Sagrada Família challenges the assumption that verticalization directly correlates with urban access and infrastructure. In this case, even with predominantly low-rise buildings, the neighborhood benefits from a high concentration of services, commercial nodes, and robust accessibility, which reinforces its centrality in the urban structure. This observation highlights that verticalization, while often used as a proxy for urban development, is not universally indicative of socio-spatial quality.

Within this cluster, several distinct urban subtypes can be identified, from highly consolidated, service-rich zones, like Sagrada Famılia, to more fragmented areas with comparable building typologies but lower infrastructural support. Finally, Figure 7D shows consolidated areas with higher verticalization, where building height correlates with a greater number of local centralities and better infrastructure distribution.

The 4x2 model was applied to help refine the classification of areas where urban patterns showed ambiguous or overlapping characteristics (Figure 8).

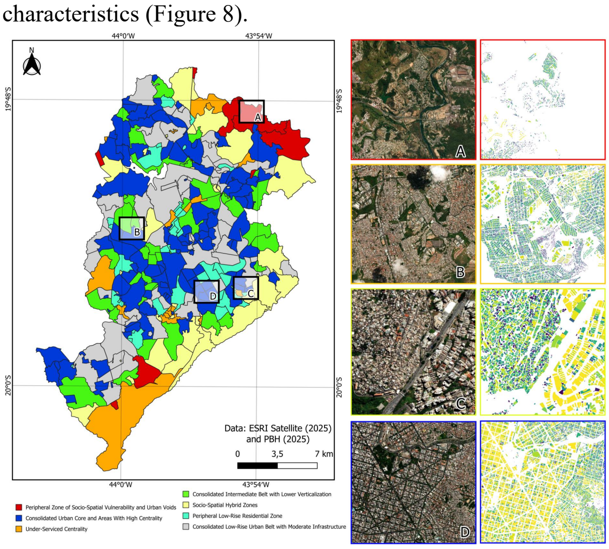

Figure 8. Cluster 4x2

In Figure 8A, we observe the neighborhood Conjunto Capitão Eduardo, located on the boundary between Belo Horizonte and the municipality of Santa Luzia. This area is marked by socio-spatial vulnerability and limited infrastructure, serving as a representative example of the red cluster associated with peripheral and marginalized urban zones.

Figure 8B, in contrast, shows the Vera Cruz neighborhood. While it also exhibits infrastructural shortcomings, it differs from 8A by presenting multiple local centralities and urban nodes, which suggest a more complex and possibly transitional socio-spatial condition, not fully aligned with the purely peripheral typology. In Figure 8C, we identify areas that remain mixed even under the 4x2 model con- figuration. These zones present challenges for classification due to the coexistence of sharply contrasting urban forms within small spatial units.

A notable example is the juxtaposition of Sion, a highly verticalized, upper-middle-class neighborhood, and Vila Barragem Santa Lúcia, an informal settlement, separated merely by a single avenue. These cases highlight the need for future studies to adopt finer, intra-neighborhood analytical scales and potentially propose new sub-clusters to capture this internal complexity.

Finally, Figure 8D depicts areas with a clearer definition of consolidated urban cores, characterized by high-rise buildings and intense verticalization, reinforcing their classification within central, formalized urban typologies.

4. Discussion

The spatial distribution of verticalization in Belo Horizonte reflects a classic pattern of urban centrality and peripheral exclusion. As demonstrated in the verticalization map (Figure 6), the highest concentration of tall buildings is situated in the city center, forming a radial gradient that diminishes progressively toward the outskirts. This is consistent with broader trends observed in several cities, where land value, infrastructure investments, and urban consolidation are typically concentrated in core areas (Puspitasari and Kwon, 2020). The use of verticalization as a proxy for urban development proves to be insufficient when considered in isolation.

The clustering analysis using Self-Organizing Maps (SOM) provides a multi-dimensional nature view of socio-spatial dynamics by integrating verticalization with infrastructure availability and demographic indicators. In the 2x2 SOM model, broad urban typologies emerge. Cluster A (Figure 7A), encompassing highly vulnerable communities such as Vitória, illustrates how the lack of vertical development is not simply a matter of urban form but a reflection of systemic exclusion from the formal real estate market.

The absence of centralities and infrastructural support reinforces a cycle of marginalization, demonstrating that low verticalization is often symptomatic of broader socio-economic precarity. In contrast, cluster B (Figure 7B) exemplifies hybrid and transitional zones such as Jardim Industrial. These areas illustrate the coexistence of formal and informal elements, both in built structure and socio-economic conditions. Cluster C (Figure 7C), represented by the Sagrada Família neighborhood, challenges simplistic correlations between building height and urban quality. Despite having low verticalization, the area benefits from a dense concentration of services and strong local centrality. This underscores that urban quality can emerge from non-vertical, well-connected, and infrastructurally sound environments.

Furthermore, the internal variation within this cluster suggests that multiple subtypes exist even within broadly defined typologies. Cluster D (Figure 8D) represents consolidated high-rise zones that align with conventional markers of urban development, showing a connection between building height, centralities, and infrastructure provision.

The 4x2 SOM model refined these distinctions by capturing ambiguous and overlapping urban patterns. For instance, in Figure 8C, the juxtaposition of Sion and Vila Barragem Santa Lucia illustrates the socio-spatial fragmentation that often occurs within a single administrative unit. These micro-contradictions point to the limitations of neighborhood-scale analyses and reinforce the need for finer spatial units in future studies.

Moreover, the presence of zones like Vera Cruz (Figure 8B), which combine infrastructural scarcity with multiple centralities, complicates typological classification and suggests the emergence of new, transitional urban forms. The SOM methodology shows effectiveness in surfacing such ambiguities, offering a flexible approach that accommodates both continuity and contrast in urban morphology.

Overall, the results reveal that while verticalization is a relevant spatial indicator, it must be interpreted within a broader socio-infrastructural framework. The combination of LiDAR-derived building height data with urban infrastructure and demographic proxies enriches the analytical landscape, enabling a more nuanced understanding of urban inequality, spatial fragmentation, and morphological processes.

Conclusion

This study employed LiDAR data and Self- Organizing Maps (SOM) to investigate patterns of verticalization and infrastructure in the city of Belo Horizonte. By integrating building height with socio-spatial variables such as local centralities and the distribution of informal settlements, the research offers a multidimensional perspective on urban morphology and inequality.

The results confirm that verticalization in Belo Horizonte follows a radial spatial logic, concentrated in central areas and gradually declining toward the periphery. However, the analysis also demonstrates that verticality alone is not a definitive marker of urban development or quality. Through unsupervised clustering, we identified a variety of urban typologies, including highly consolidated zones, hybrid and transitional areas, and socio-spatially vulnerable regions.

The use of a dual SOM configuration (2x2 and 4x2) allowed for both general and detailed classifications, revealing the complexity of intra-urban dynamics. Important were the findings in mixed zones where formal and informal structures coexist in close proximity, exposing the need for finer spatial analyses and more granular classification schemes.

This methodological framework, combining LiDAR, spatial overlays, and machine learning, proves to be effective in detecting latent patterns and contradictions in urban space. The approach can be replicated in other cities and extended with additional variables such as land use, accessibility indices, or temporal changes in verticalization. Future research should focus on temporal analysis to assess the evolution of verticalization and socio-spatial fragmentation. Additionally, exploring causality between urban form and social outcomes may enrich planning decisions aimed at reducing inequality and fostering inclusive development.

Acknowledgements

The authors acknowledge the financial support from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil (Finance Code 001); FAPESP (Grant 2020/09215-3); CNPq (Grant 311324/2021-5); SAC (Grant 00413/2025); the Postgraduate Program in Remote Sensing (PGSER) at the National Institute for Space Research (INPE) and the PostGraduate Program in Aeronautical Infrastructure Engineering (PG-EIA) at the Aeronautics Institute of Technology. The authors are also grateful for the institutional support provided by the PGSER-INPE.

References

Bonczak, B., Kontokosta, C. E., 2019. Large-scale parameterization of 3D building morphology in complex urban landscapes using aerial LiDAR and city administrative data. *Computers, Environment and Urban Systems,* 73, 126–142.

Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S., Çöltekin, A., 2015. Applications of 3D City Models: State of the Art Review. *ISPRS International Journal of Geo-Information*, 4(4), 2842-2889.

Dewez, T. J. B., Girardeau-Montaut, D., Allanic, C., Rohmer, J., 2016. Facets: a cloudcompare plugin to extract geological planes from unstructured 3D point clouds. *ISPRS Archives*, XLI-B5, 799–804.

Ding, Q. et al., 2018. 3d rock slope data acquisition by photogrammetry approach and extraction of geological planes using facet plugin in cloudcompare. *IOP Conference Series: Earth and Environmental Science*, 169number 1, 012051.

Goovaerts, P., 2019. Kriging Interpolation. *Geographic Information Science & Technology Body of Knowledge*, 2019(Q4).

Harris, A., 2015. Vertical urbanisms. *Progress in Human Geography*, 39(5), 601–620.

Inostroza, L., 2017. Informal urban development in Latin American urban peripheries. Spatial assessment in Bogota´, Lima and Santiago de Chile. *Landscape and Urban Planning*, 165, 267–279.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43(1), 59–69.

Lao, J. et al., 2021. Retrieving building height in urban areas using ICESat-2 photon-counting LiDAR data. *International Journal of Applied Earth Observation and Geoinformation*, 104, 102596.

Li, G., Li, F., 2019. Urban sprawl in China: Differences and socioeconomic drivers. *Science of The Total Environment*, 673, 367–377.

Ma, X. et al., 2023. Mapping fine-scale building heights in urban agglomeration with spaceborne lidar. *Remote Sensing of Environment*, 285, 113392.

Mahtta, R., Mahendra, A., Seto, K. C., 2019. Building up or spreading out? Typologies of urban growth across 478 cities of 1 million+. *Environmental Research Letters*, 14(12), 124077.

Prefeitura de Belo Horizonte, 2025a. ´Indice de desenvolvimento humano municipal (idh-m). Accessed: Jun. 12, 2025.

Prefeitura de Belo Horizonte, 2025b. Acervo Cartográfico e Geográfico. Accessed: Jun. 12, 2025.

Prefeitura de Belo Horizonte, 2025c. BH Map- Visualizador. Accessed: Jun. 12, 2025.

Puspitasari, A. W., Kwon, J., 2020. Comparison of spatial layout of tall buildings clustered in circular, rectangular, and linear geographical areas and impact on skyline. *Buildings*, 10(4), 64.

Rajendra, Y. D. et al., 2014. Evaluation of partially overlapping 3D point cloud's registration by using ICP variant and cloudcompare. *ISPRS Archives*, XL–8(1), 891–897.

Shi, W., Meng, Q., Zhang, L., Wang, J., Zhou, T., Atkinson, P. M., 2024. Building height extraction from monocular off-nadir satellite sensor imagery. *Proc. IGARSS*, 2537–2540.

Wang, N., Chen, Z., Li, T., Zhen, M., 2022. Spatiotemporal Pattern Evolution and Influence Mechanism of Urban Vertical Expansion: A Case Study of Jiangsu Province, China. *Land*, 11(3), 433.

Zambon, I., Colantoni, A., Salvati, L., 2019. Horizontal vs vertical growth: Understanding latent patterns of urban expansion in large metropolitan regions. *Science of The Total Environment*, 654, 778–785.