AI-BASED BUILDING INSTANCE SEGMENTATION IN FORMAL AND INFORMAL SETTLEMENTS

Philipp Schuegraf[1], Dorothee Stiller[2], Jiaojiao Tian[1], Thomas Stark[2], Michael Wurm[2], Hannes Taubenböck[2,3], Ksenia Bittner[1]

[1]German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF)
[2]German Aerospace Center (DLR), German Remote Sensing Data Center (DFD)
[3]Department of Urban Remote Sensing, Institute of Geography and Geology, University of Würzburg

ABSTRACT

Building instances play a pivotal role in understanding population distribution and assessing vulnerability in the face of potential risks. Building instance segmentation is a valuable technique for identifying individual structures, however, complex urban environments pose great challenges, especially in the informal areas. In this study, we utilize a building instance segmentation method specifically designed to discern single buildings in both formal and informal settlements. Employing a SkipFuse-UResNet34 model, we generate building instances for Medellín, Colombia, resulting in a more comprehensive building mask compared to conventional official data sources. This enhanced mask serves as a vital tool for estimating the population at risk, enabling a thorough comparison with official data and addressing current spatial knowledge gaps.

Index Terms— building instance segmentation, informal settlements, deep learning, urbanization, satellite imagery

1. INTRODUCTION

In the undulating slopes of Medellín, Colombia, where challenging topography intertwines with the vulnerability of communities, the threat of landslides emphasizes the pursuit of precise insights into this complex urban landscape. We urgently require additional information to assess the potential risks and determine the necessity of employing AI for the instance segmentation of buildings. Medellín, faces unique challenges as it is nestled in a valley surrounded by steep slopes. Historical urbanization, driven by rapid industrial and economic growth in the mid-20th century, led to an influx of migrants settling informally on the city's outskirts [1]. The escalation of informal housing, exacerbated by conflicts between paramilitary forces and guerrilla groups, particularly in later years, has extended the city into precarious, hard-to-reach areas on these steep slopes.

The informal settlements, characterized by low-quality building fabric, are highly vulnerable to landslide hazards due to frequent heavy rainfall and the presence of weak, erosive rocks in the bedrock [2]. Effectively countering this risk requires precise knowledge of the at-risk areas and an understanding of the potential impact on the population. However, official population data for Medellín accurately geolocates formal residents but significantly underrepresents the more recent informal settlements on steep slopes, leading to a substantial underestimation of the exposed population [3].

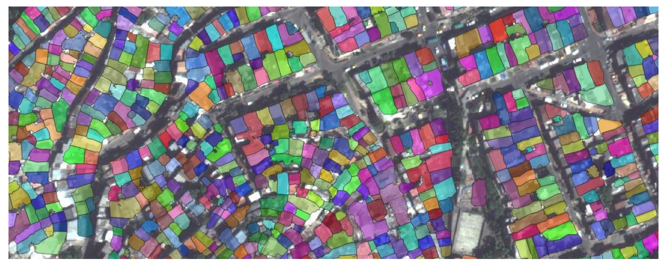


Fig. 1: Example of selected test area in Medellin and the obtained instance segmentation results from the proposed methodology.

Detecting informal settlements poses significant challenges [4, 5] due to limited data availability and the uncertainty associated with reference data [6, 7]. This difficulty is further aggravated by the intra-and-inter urban variability observed in informal areas [8]. While the utilization of very high-resolution imagery proves advantageous for informal settlement classification, as demonstrated in [9], the detection of individual buildings within informal settlements remains an extremely challenging task, often requiring manual intervention [10].

To address this critical knowledge gap, we leverage high-resolution remote sensing imagery in combination with a digital surface model (DSM), capable of detecting single buildings in the challenging urban environment of Medellín. This approach will be tested in formal settlements, but even more pertinent and demanding is its application in the city's informal settlements (see Fig. 1). The morphological characteristics of these areas, marked by small-scaled, intricate, and

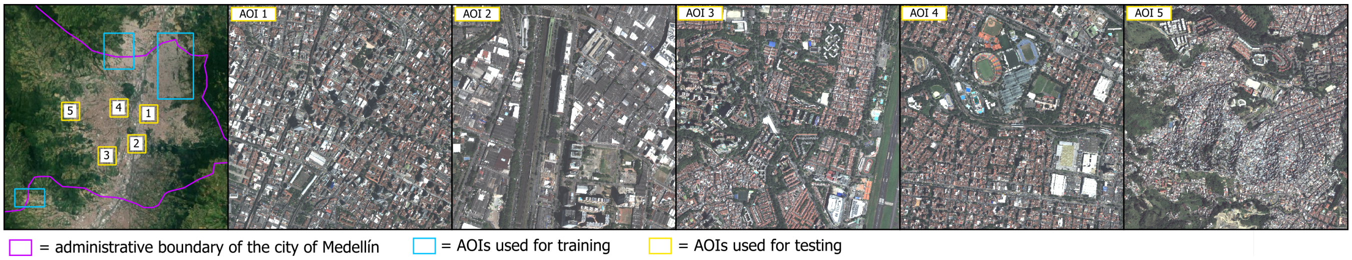


Fig. 2: Locations of the AOIs within the administrative area of Medellín and detailed maps for the AOIs used for testing.

densely packed structures, are indicative of informal settlements [11]. In this study, we employ a SkipFuse-UResNet34 in order to generate building instances, thus leading to a more comprehensive building mask compared to official data sources. This enhanced building mask serves as a crucial tool to estimate the population at risk of landslides, allowing for a comparison with official data and filling the current spatial knowledge void.

2. STUDY AREA AND DATA

The study area is the municipality of Medellín (Fig. 2). Three areas of interest (AOI) were used for training, each with different sizes, i.e. from east to west, 3x5.5km, 2.5x3km, and 2.5x1.5km. The testing was undertaken in five distinct AOIs within Medellín, each covering an expanse of 1.5x1.5km. The selection of the test AOIs was guided by two primary considerations. Firstly, they were chosen for their high building density, rendering them well-suited for our proposed application. Secondly, these AOIs showcase a diverse array of morphological building types, effectively capturing the inherent structural complexity of buildings. AOI 1 is situated in the central business district (CBD), AOI 2 is characterized by a substantial industrial area, AOI 3 primarily consists of residential building structures with varying heights, AOI 4 incorporates large communal and industrial buildings alongside low- and mid-rise structures, and AOI 5 features a densely-built informal area located at the steep slopes on the outskirts of Medellín.

For the study area of Medellín, we use very high resolution RGB satellite imagery from WorldView-3 with a geometric resolution of $0.3\,\mathrm{m}$. For training the model, additional imagery data depicting Berlin city from WorldView-4, Bonn city from open source aerial data, and Hamburg city from WorldView-2 were used. All images were brought to a geometric resolution of $0.3\,\mathrm{m}$.

In contrast to prior research, our model utilized a pansharpened RGB image in addition to DSM generated by semi-global matching (SGM) technique [12], rather than relying on a single panchromatic channel and DSM. The RGB images enables better visualization and interpretation of the details present in the scene. For a challenging task like detecting

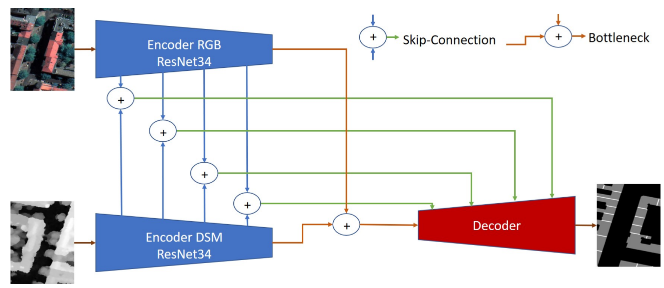


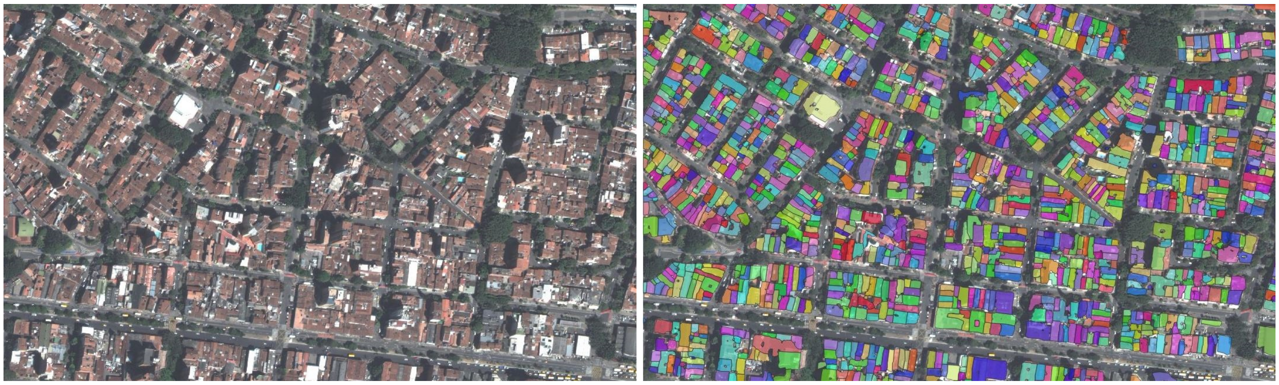
Fig. 3: Visualization of the utilized network architecture.

small buildings in informal areas, spectral information plays a crucial role. It provides the network with additional clues about various structures, particularly in cases of dissimilar textures.

We rely on official building cadaster data for training, validation, and testing of our method. Cadaster data from Germany for the cities of Berlin, Bonn, and Hamburg, were used for training and validation. In addition to these, cadaster data from Medellín was also used for the purpose of training, validation, and testing. This dataset encompasses the footprints for a significant portion of building structures in Medellín. Nevertheless, it is important to highlight that the official building cadaster does not include numerous informal or recently constructed buildings.

3. METHODOLOGY

With this paper, we are testing the applicability of our previously developed methodology [13] for building sections instance segmentation on a new challenging area of Medellín. The proposed method comprises two consecutive steps: Initially, a SkipFuse-UResNet34 architecture (see Figure 3) is used to segment buildings and separation lines between building sections as a 3-class problem. The process utilizes RGB and DSM images fed into two distinct encoders. To preserve fine-grained spatial information, feature maps, obtained at four distinct scales from the two encoders, are aggregated by summation and serve as the input for the full-scale skip-connections. During the network training, a combination of three losses—Weighted Cross Entropy Loss, Dice Loss, and



(a) Sample of area from AOI 4 showing a formal (city) region (b) Instance segmentation results for sample of area from AOI 4

(c) Sample of area from AOI 5 showing an informal settlement (d) Instance segmentation results for sample of area from AOI 5

Fig. 4: Detailed visual analysis of building instance segmentation results on two challenging areas from AOI 4 and 5.

Topology Loss—was employed. For more details, please, refer to [13].

Next, a map of building section instances is generated using the watershed transform [14] as a post-processing step. Mainly, the watershed transform interprets the obtained 3-class map of background, building, and separation line together with a seed image and a mask as a topographical surface. The seed map and mask are extracted from the predicted information about buildings and separation lines. Subsequently, watershed transform simulates a flooding scenario where water initiates flooding from the seeds and settles into basins. These basins are marked by watershed lines, aligning with high image intensities. The mask confines the virtual water flow to specific regions. The enclosed regions delineated by watershed lines are then recognized as objects.

We conduct evaluations on five selected AOIs that combine both formal urban environments and informal settlements. Those complex areas are selected for the purpose to better demonstrate the strength of the proposed methodology.

4. RESULTS AND DISCUSSION

Figure 4 illustrates two samples of complex and very dissimilar regions of Medellín together with the obtained instance segmentation results from our proposed methodology. In a city region (see Fig. 4a), buildings can have very complex shapes with many sections within one building construction. To distinguish those sections or separation lines between those sections is a challenging task even for a human eye. Nevertheless, one can recognize a city layout within this region, which cannot be done by observing the informal settlement area (Fig. 4c). Although buildings within informal settlements often exhibit simple rectangular shapes, their dense arrangement makes it challenging to distinguish each individual construction. Analysing the obtained instance segmentation results for both regions in Fig. 4b and Fig. 4d, we can say

Table 1: Quantitative results for IOU, FPR, FNR metrics of building class and overall accuracy evaluated on five selected AOIs for testing.

AOI	IoU_{BLD}	FPR_{BLD}	FNR_{BLD}	OA
1	0.694	0.125	0.234	0.816
2	0.761	0.039	0.192	0.902
3	0.669	0.066	0.223	0.889
4	0.708	0.078	0.198	0.878
5	0.620	0.086	0.254	0.866

that our developed methodology enables the identification of even the smallest building segments, regardless of the building complexity or the region.

To quantify the quality of instance segmentation results, we evaluate the metrics intersection over union (IOU), false positive rate (FPR), false negative rate (FNR) and overall accuracy (OA) and report their performance in Table 1. Our approach achieves an IoU around or above 0.7 on the AOIs 1-4. In the informal AOI 5, it still scores 0.62 IoU, whereas the overall accuracy (OA) is high at above 0.8 for all AOIs. The false positive rate (FPR) is very low for all AOIs, which indicates that the method produces few false positives. On the other hand, the resulting false negative rate (FNR) is above 0.19 for all AOIs, pointing at the issue that many small buildings are very hard to detect.

5. CONCLUSION

Our proposed building instance segmentation approach is able to identify single building instances in both settlement types, formal and informal. This framework can obtain detailed instance segmentation masks, especially for informal regions, facilitates the accurate counting of built houses and estimation of the population residing in these areas. This information becomes critical during disasters and for coordinating humanitarian aid efforts.

6. REFERENCES

[1] John J Betancur, "Approaches to the regularization of informal settlements: the case of primed in medellin, colombia," *Global Urban Development*, vol. 3, no. 1, pp. 1–15, 2007.

[2] Joseph Claghorn and Christian Werthmann, "Shifting ground: Landslide risk mitigation through community-based landscape interventions," *Journal of Landscape Architecture*, vol. 10, no. 1, pp. 6–15, 2015.

[3] Marta Sapena, Marlene Kühnl, Michael Wurm, Jorge E Patino, Juan C Duque, and Hannes Taubenböck, "Empiric recommendations for population disaggregation under different data scenarios," *Plos one*, vol. 17, no. 9, pp. e0274504, 2022.

[4] Monika Kuffer, Karin Pfeffer, and Richard Sliuzas, "Slums from space—15 years of slum mapping using remote sensing," *Remote Sensing*, vol. 8, no. 6, pp. 455, 2016.

[5] Michael Wurm, Thomas Stark, Xiao Xiang Zhu, Matthias Weigand, and Hannes Taubenböck, "Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 150, pp. 59–69, 2019.

[6] Nicolas Johannes Kraff, Michael Wurm, and Hannes Taubenböck, "Uncertainties of human perception in visual image interpretation in complex urban environments," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 4229–4241, 2020.

[7] Thomas Stark, Michael Wurm, Xiao Xiang Zhu, and Hannes Taubenböck, "Detecting challenging urban environments using a few-shot meta-learning approach," in *2023 Joint Urban Remote Sensing Event (JURSE)*. IEEE, 2023, pp. 1–4.

[8] Hannes Taubenböck, Nicolas J Kraff, and Michael Wurm, "The morphology of the arrival city-a global categorization based on literature surveys and remotely sensed data," *Applied Geography*, vol. 92, pp. 150–167, 2018.

[9] Michael Wurm, Raphael Tubbesing, Thomas Stark, Marlene Kühnl, Marta Sapena, Wolfgang Sulzer, and Hannes Taubenböck, "Revealing landslide exposure of informal settlements in medellín using deep learning," in *2023 Joint Urban Remote Sensing Event (JURSE)*. IEEE, 2023, pp. 1–4.

[10] Nicolas J Kraff, Hannes Taubenböck, and Michael Wurm, "How dynamic are slums? eo-based assessment of kibera's morphologic transformation," in *2019 Joint Urban Remote Sensing Event (JURSE)*. IEEE, 2019, pp. 1–4.

[11] Michael Wurm and Hannes Taubenböck, "Detecting social groups from space–assessment of remote sensing-based mapped morphological slums using income data," *Remote Sensing Letters*, vol. 9, no. 1, pp. 41–50, 2018.

[12] Pablo d'Angelo and Peter Reinartz, "Semiglobal matching results on the isprs stereo matching benchmark," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, pp. 79–84, 2012.

[13] Philipp Schuegraf, Stefano Zorzi, Friedrich Fraundorfer, and Ksenia Bittner, "Deep learning for the automatic division of building constructions into sections on remote sensing images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023.

[14] Serge Beucher and Fernand Meyer, "The morphological approach to segmentation: the watershed transformation," in *Mathematical morphology in image processing*, pp. 433–481. CRC Press, 2018.