The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIX-B2-2026
XXV ISPRS Congress 2026 "From Imagery to Understanding", Commission II, 4–11 July 2026, Toronto, Canada

Augmenting City Models with Handheld LiDAR and 3D Gaussian Splatting for Inclusive Pedestrian Infrastructure Assessment

Deni Suwardhi [1,2], Wahyunan Andika [2], Ratri Widyastuti [1], Widiaomoko Azis Fadilah [3], Arnadi Murtiyoso [3], Pierre Grusennmeyer [3], Fabio Remondino [4], Farhan Helmy [5]

[1] Spatial System and Cadastral Research Group, Institut Teknologi Bandung (ITB), Indonesia – (deni.suwardhi, ratri)@itb.ac.id
[2] PT Inovasi Mandiri Pratama, Spatial Information Company, Indonesia - wahyunan.andika@inovamap.com
[3] Université de Strasbourg, CNRS, INSA Strasbourg, ICube Laboratory UMR 7357, Photogrammetry and Geomatics Group, 67000, Strasbourg, France– (widiatmoko-azis.fadilah, arnadi.murtiyoso, pierre.grussenmeyer)@insa-strasbourg.fr
[4] 3D Optical Metrology (3DOM) Unit, Bruno Kessler Foundation (FBK), Trento, Italy – remondino@fbk.eu
[5] Advanced System Computing, Design and Innovation (ASCODI) Laboratory, Indonesia – farhan.helmy@dilansindonesia.org

Keywords: Handheld lidar, 3D Gaussian Splatting, CityGML, Digital Twin, Pedestrian Accessibility, Urban Modelling.

Abstract

Urban digital twins increasingly require pedestrian-scale three-dimensional (3D) representations to support accessibility and inclusiveness assessment. However, existing approaches typically emphasize either geometric accuracy or visual realism, while lacking an integrated framework for analysing pedestrian-level conditions. This study proposes a hybrid workflow integrating handheld LiDAR and 3D Gaussian Splatting (3DGS) within a CityGML-based semantic framework for accessibility assessment. Handheld LiDAR provides centimetre-level geometric measurements, enabling the extraction of key indicators such as slope, surface roughness, and obstacle presence. In parallel, 3DGS reconstruction from 360° video imagery enhances visual realism and perceptual understanding. Both datasets are co-registered and structured within the CityGML 3.0 Transportation model to represent pedestrian environments in a unified spatial and semantic framework. Accessibility assessment was conducted using three approaches: LiDAR-based analysis, field survey observations, and immersive evaluation in a Virtual Reality (VR) environment. The LiDAR-based results were used as reference. Comparative analysis shows that field survey assessment achieves an agreement of approximately 85.7%, while VR-based assessment reaches approximately 75.4%. The results indicate that while VR does not replace metric-based analysis, it enables perception-driven and participatory evaluation. In particular, VR-based assessment shows potential to involve users, including people with disabilities, in accessibility evaluation through immersive and remote interaction. The proposed approach contributes to the development of human-scale urban digital twins by integrating metric accuracy, semantic structure, and participatory evaluation for more inclusive accessibility analysis.

1. Introduction

The development of urban digital twins increasingly relies on accurate and semantically rich three-dimensional (3D) geospatial data representations to support sustainable, inclusive, and data-driven decision-making (Scolamiero et al., 2025). In this context, pedestrian environments, including sidewalks, crosswalks, and ramps, play a critical role in ensuring accessibility and mobility for all citizens, especially people with disabilities and the elderly. In Southeast Asian developing cities, pedestrian environments are particularly complex due to the coexistence of diverse street activities, mixed land uses, and competing demands between pedestrians and vehicles (Miura et al., 2024).

While 3D city models such as CityGML have been widely adopted for urban representation, they are primarily focused on buildings and large-scale structures, and often lack the geometric detail and semantic granularity required for pedestrian-level accessibility assessment (Beil et al., 2020). As a result, critical aspects such as surface continuity, curb height, and small-scale obstacles are not adequately represented within existing urban digital twin frameworks.

Recent advances in image-based 3D reconstruction, such as Neural Radiance Fields (NeRF) and 3D Gaussian Splatting (3DGS) (Kerbl et al., 2023), have enabled photorealistic visualization of complex urban environments. However, these approaches remain limited in geometric accuracy and structural reliability, particularly in outdoor pedestrian contexts. In contrast, LiDAR-based methods provide centimetre-level geometric precision but lack realistic visual representation, which is essential for perception-based evaluation and participatory analysis.

This trade-off between geometric accuracy and visual realism has motivated the development of multimodal approaches that integrate LiDAR and radiance-field representations. Recent studies such as GauU-Scene V2 (Xiong et al., 2024) and LI-GS (Jiang et al., 2025) demonstrate the potential of combining these data sources for large-scale reconstruction. However, these approaches primarily operate at the city or district scale and do not explicitly address pedestrian-level accessibility assessment or human-scale evaluation.

Figure 1 illustrates the conceptual framework of the proposed pedestrian-scale urban digital twin, highlighting the integration of geometric, visual, and semantic components for accessibility assessment. The framework combines handheld LiDAR for accurate geometric measurement, 3DGS for photorealistic representation, and CityGML-based semantic modelling to structure pedestrian elements such as sidewalks, crossings, and street furniture. In addition, the framework incorporates multi-source evaluation, including LiDAR-based analysis, field observations, and immersive VR-based assessment, enabling both quantitative and perception-driven evaluation of pedestrian infrastructure. This integrated approach positions the proposed method as a comprehensive solution for inclusive urban accessibility analysis.

Despite these advances, existing studies rarely integrate semantic modelling and multi-perspective validation for pedestrian accessibility assessment. In particular, the combination of metric measurements, real-world observations, and perception-based evaluation within a unified digital twin framework remains limited, especially in complex and obstacle-rich pedestrian environments.

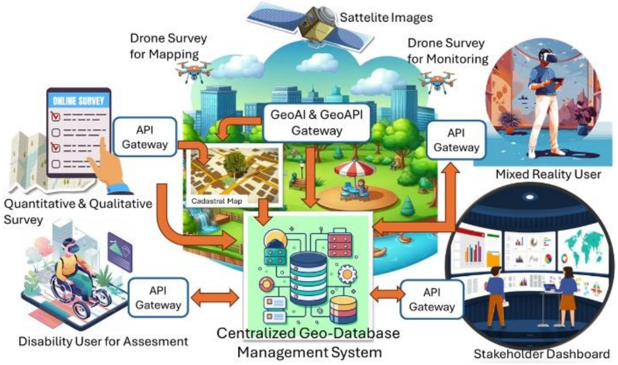

Figure 1. Digital twin framework for the monitoring of city infrastructure’s inclusiveness.

This study aims to develop a hybrid framework for pedestrian-scale accessibility assessment by integrating handheld LiDAR and 3DGS within a CityGML-based semantic digital twin. The proposed approach combines metric geometry, photorealistic representation, and semantic modelling to support comprehensive analysis of pedestrian infrastructure.

The main contribution of this study lies in the integration of metric, semantic, and perceptual representations within a unified pedestrian-scale digital twin, supported by a multi-source accessibility assessment framework.

2. Data and Methodology

The proposed methodology is designed to support pedestrian-scale accessibility assessment within a multi-layer urban digital twin framework. It integrates geometric accuracy, photorealistic representation, and semantic modelling to enable both quantitative and perception-driven analysis of pedestrian infrastructure. The methodology consists of four main components: (i) multimodal data acquisition using handheld LiDAR and 360° video imagery, (ii) geometric and radiance-field reconstruction, (iii) semantic modelling based on CityGML, and (iv) multi-source accessibility assessment combining metric, observational, and immersive evaluation.

The study was conducted in Sumur Bandung District, Indonesia, a historic urban area characterized by narrow sidewalks, frequent obstacles, and mixed pedestrian–vehicle interactions. These conditions make it a representative environment for evaluating pedestrian accessibility at a human scale.

Figure 2 illustrates the overall workflow of the proposed framework, which integrates geometric, visual, and semantic components into a unified pedestrian-scale digital twin. The workflow establishes a relationship between data acquisition, reconstruction, and assessment, enabling both quantitative analysis and perception-driven evaluation of pedestrian infrastructure conditions.

The acquired datasets are processed to generate three complementary representations of the pedestrian environment. The LiDAR point cloud is used to derive geometric information and to construct a semantically structured model based on CityGML, representing pedestrian-related objects such as sidewalks and street furniture. In parallel, image-based reconstruction is used to generate a photorealistic representation through 3DGS, which is co-registered with the LiDAR dataset. These representations are integrated within an immersive visualization environment, enabling combined geometric, semantic, and perceptual analysis.

Pedestrian accessibility is then evaluated using three complementary approaches: (i) LiDAR-based geometric analysis, (ii) field survey observations, and (iii) immersive assessment within a Virtual Reality (VR) environment. The results from these approaches are compared to support comprehensive and multi-perspective evaluation of pedestrian infrastructure conditions.

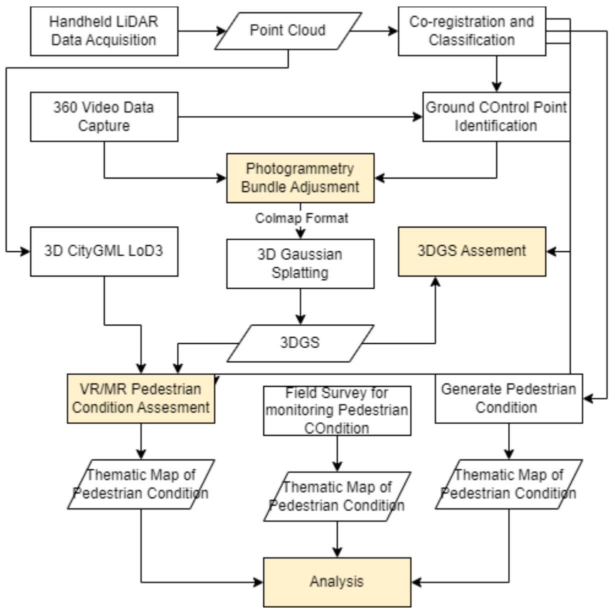

Figure 2. Hybrid workflow integrating handheld lidar, 3DGS, and CityGML semantics.

2.1 Data Acquisition

The pedestrian environment in the study area was documented using a combination of handheld LiDAR scanning and 360° video capture. The LiDAR system was used to acquire accurate geometric information of the sidewalk surface and surrounding urban objects, while the 360° video recordings provided detailed visual information of the pedestrian environment. This multimodal acquisition approach enables the combined documentation of metric geometry and photorealistic appearance required for pedestrian accessibility assessment.

Geometric data were acquired using a Geosun GS-120G handheld LiDAR system. The sensor was operated along pedestrian pathways following a walking trajectory to capture the surface geometry of sidewalks and surrounding street elements. During the acquisition, the operator walked along the sidewalk corridor to ensure continuous coverage of the pedestrian environment. The handheld LiDAR system generates dense point clouds with centimeter-level accuracy, enabling detailed representation of sidewalk surfaces and surrounding urban objects such as street furniture, trees, and lighting poles. The resulting point cloud dataset serves as the primary geometric reference for subsequent processing and analysis.

Visual data were collected using a GoPro Max 360° camera mounted on a helmet worn by an operator walking along the same pedestrian corridor. The camera records omnidirectional video capturing the visual appearance of the pedestrian environment, including surface textures, lighting conditions, and surrounding urban objects. The recording was conducted along the entire sidewalk segment to ensure continuous visual coverage of the study area. These video sequences serve as the input data for subsequent frame extraction and photogrammetric reconstruction processes.

In addition to point cloud acquisition, the handheld LiDAR system also provides synchronized fisheye imagery, which can be utilized as an additional visual source for image-based reconstruction and radiance-field modelling. This additional dataset enables comparative evaluation of different image sources in subsequent processing stages.

2.2 360° Video Pre-processing

The recorded 360° video was processed to generate input images suitable for photogrammetric reconstruction. Frames were sampled at a reduced rate to maintain consistency with available metadata while minimizing data redundancy and computational cost.

Since the extracted frames are stored in an equirectangular projection, they were converted into multiple perspective images to ensure compatibility with standard photogrammetric processing. This transformation improves feature visibility and overlap between adjacent views, enabling reliable feature matching and camera pose estimation.

The associated metadata were then assigned to the generated images to produce geotagged photographs, preserving their spatial relationship within the original acquisition trajectory. The resulting image set was used as input for subsequent photogrammetric reconstruction, providing a geometrically consistent and efficient dataset for further processing.

2.3 Photogrammetric Reconstruction

The geotagged perspective images generated from the 360° video pre-processing stage were imported into Agisoft Metashape for photogrammetric reconstruction. An initial image alignment process was performed to estimate camera poses and generate a sparse point cloud based on feature matching and bundle adjustment. This initial stage produces a geometrically consistent relative reconstruction of the scene based solely on image content.

Following the alignment stage, the transformation parameters of the reconstruction were reset to remove the influence of low-accuracy positional metadata derived from the video. This step ensures that the reconstruction is treated as a purely relative model prior to georeferencing, avoiding bias introduced by potentially inaccurate trajectory data.

To establish an accurate spatial reference, control points derived from the handheld LiDAR dataset were introduced into the photogrammetric model. These control points correspond to identifiable features visible in both the images and the LiDAR point cloud, enabling precise spatial linkage between the image-based reconstruction and the LiDAR geometry. The LiDAR dataset serves as the primary geometric reference due to its higher metric accuracy and stability compared to trajectory-based positioning.

The control points were incorporated into a bundle adjustment optimization process to refine camera parameters and improve geometric consistency. This optimization minimizes reprojection errors while enforcing alignment with the LiDAR-derived control points, thereby reducing scale ambiguity and drift effects commonly encountered in image-based reconstruction.

As a result, the reconstructed model is transformed from a relative coordinate system into a georeferenced model within the Universal Transverse Mercator (UTM) coordinate system (Zone 48S). This georeferencing ensures spatial consistency with the LiDAR dataset and enables seamless integration with subsequent representations, including 3DGS and CityGML-based semantic models.

This step is critical for the proposed framework, as it establishes a metrically reliable bridge between image-based and LiDAR-based representations. By anchoring the photogrammetric reconstruction to the LiDAR reference, the workflow ensures that subsequent analyses, including geometric measurement and immersive evaluation, are performed within a consistent and accurate spatial framework.

2.4 3D Gaussian Splatting Reconstruction

The georeferenced datasets obtained from photogrammetric reconstruction and LiDAR acquisition were used as input for 3D Gaussian Splatting (3DGS) modelling. Two types of input data were considered: (i) images derived from 360° video processed through photogrammetry, and (ii) fisheye images associated with the handheld LiDAR system.

The use of these two datasets enables a comparative evaluation of how different acquisition modalities influence the geometric consistency and visual quality of the resulting 3DGS models. While the GoPro-based dataset provides broader visual coverage, the LiDAR-based imagery is inherently linked to a more stable trajectory, potentially improving spatial consistency.

Prior to training, a coordinate normalization step was applied to shift the georeferenced data into a local coordinate system, improving numerical stability during processing. The datasets were then exported in COLMAP format, including camera parameters and image information, as input for the 3DGS pipeline.

The 3DGS models were generated using the Postshot Splat3 implementation, which enables efficient photorealistic reconstruction suitable for immersive visualization (Fadilah et al., 2026). The resulting models provide visually continuous representations of the pedestrian environment and are spatially aligned with the LiDAR point cloud.

This step enables the integration of metric and photorealistic representations within a unified spatial reference system and supports subsequent analysis of geometric accuracy and perception-driven assessment.

2.5 LiDAR Classification and Semantic Modelling

The georeferenced LiDAR point cloud obtained from the handheld scanning process was used as the primary dataset for geometric analysis and semantic modelling. Due to its centimeter-level accuracy, the LiDAR data provide a reliable representation of the physical characteristics of the pedestrian environment.

The point cloud was classified into pedestrian-related object categories, including sidewalks, street furniture, trees, lighting poles, bollards, and benches. This classification enables the identification of relevant elements required for pedestrian accessibility analysis and supports the transition from purely geometric data to semantically structured information.

Following classification, the segmented point cloud elements were directly mapped into a CityGML 3.0 representation at Level of Detail 3 (LoD3). Pedestrian elements were associated with appropriate semantic classes within the CityGML schema to ensure a consistent and structured representation.

This transformation enables the integration of pedestrian infrastructure into a city-scale digital twin framework while preserving both geometric and functional characteristics. The resulting semantic model complements the LiDAR-derived geometry and the 3DGS photorealistic representation, forming a multi-layer dataset for subsequent analysis.

2.6 Immersive Visualization Environment

The multi-layered datasets, including the georeferenced LiDAR point cloud, the 3DGS model, and the CityGML LoD3 semantic representation, were integrated into an immersive visualization environment to enable interactive exploration of pedestrian infrastructure.

A multi-layer approach was adopted, maintaining geometric, semantic, and photorealistic data as distinct but spatially aligned layers within a unified coordinate system (Suwardhi et al., 2026). This preserves the metric reliability of the LiDAR data, the semantic structure of the CityGML model, and the visual realism of the 3DGS representation.

The datasets are accessed through an immersive Virtual Reality (VR) environment, where users can explore the pedestrian space and switch between data layers. Visual exploration is primarily supported by the 3DGS model, semantic interpretation by CityGML, and geometric measurements by the LiDAR point cloud.

The immersive environment is designed as a virtual assessment platform, allowing users to simulate real-world pedestrian experiences, inspect sidewalk conditions, and evaluate accessibility indicators such as surface continuity, slope, and clearance from a human-scale perspective. This integration supports both visualization and perception-driven accessibility assessment.

2.7 Pedestrian Accessibility Assessment

Pedestrian accessibility assessment was conducted using three complementary approaches: (i) LiDAR-based geometric analysis, (ii) field survey observations, and (iii) immersive evaluation within a Virtual Reality (VR) environment. This multi-source strategy integrates metric measurement, real-world observation, and perception-driven evaluation within a unified framework.

A set of standardized accessibility indicators was defined, including sidewalk width, surface continuity, slope, presence of obstacles, and curb height. Based on these indicators, pedestrian segments were classified into three accessibility levels: accessible, partially accessible, and non-accessible. The LiDAR-based classification was used as the reference due to its higher geometric reliability.

Geometric indicators were quantitatively derived from the LiDAR point cloud, enabling the detection of small-scale variations such as uneven surfaces and abrupt height transitions. Field surveys were conducted along the same pedestrian corridor to document actual conditions and contextual factors, including temporary obstacles and surface degradation.

The VR-based assessment was performed within the immersive environment using the integrated multi-layer datasets. Users evaluated accessibility conditions from a human-scale perspective, supported by the photorealistic representation provided by 3DGS.

To evaluate consistency, a comparative analysis was conducted by comparing classification results across corresponding segments, with overall accuracy computed relative to the LiDAR-based reference. This framework enables a comprehensive understanding of pedestrian accessibility by combining quantitative measurement, real-world observation, and perception-based interpretation.

3. Results and Discussion

The results of the proposed hybrid framework demonstrate the effective integration of geometric accuracy, photorealistic representation, and semantic modelling for pedestrian-scale accessibility assessment. The evaluation focuses on both the geometric consistency of the multi-source datasets and the reliability of accessibility assessment derived from LiDAR-based analysis, field observations, and immersive VR-based evaluation.

The assessment was conducted based on standardized accessibility indicators defined in Section 2.7, including sidewalk width, surface continuity, slope, presence of obstacles, and curb height. These indicators were used to classify pedestrian segments into three accessibility levels: accessible, partially accessible, and non-accessible, enabling consistent comparison across different assessment approaches.

The hybrid LiDAR–3DGS workflow successfully enhances both the geometric accuracy and visual realism of the pedestrian environment. The LiDAR-derived point cloud provides centimeter-level geometric precision (RMSE ≈ 2–3 cm), while 3DGS reconstruction from 360° video imagery contributes photorealistic color and lighting, preserving fine details such as tactile paving patterns and curb transitions. The co-registration process was performed using a 3D similarity transformation based on control points derived from the LiDAR dataset, ensuring that both datasets share a unified spatial reference system and enabling direct integration with the CityGML-based semantic model.

Figure 3 illustrates the integrated multi-layer representation combining the LiDAR point cloud, 3DGS reconstruction, and CityGML-based semantic model. The visualization highlights the complementary characteristics of each representation: the LiDAR data provide accurate geometric measurements, the 3DGS model enhances visual realism and environmental perception, and the CityGML model structures pedestrian-related elements into a semantic framework. This integration enables both quantitative analysis of geometric indicators and qualitative interpretation of environmental conditions within a unified digital twin environment.

The integrated representation supports accessibility assessment by linking geometric measurements with semantic and visual information. Geometric indicators such as slope, surface continuity, and curb height are derived from the LiDAR data, while the 3DGS representation enhances the perception of surface conditions and contextual elements. The CityGML model provides a structured basis for organizing pedestrian features and supporting thematic classification.

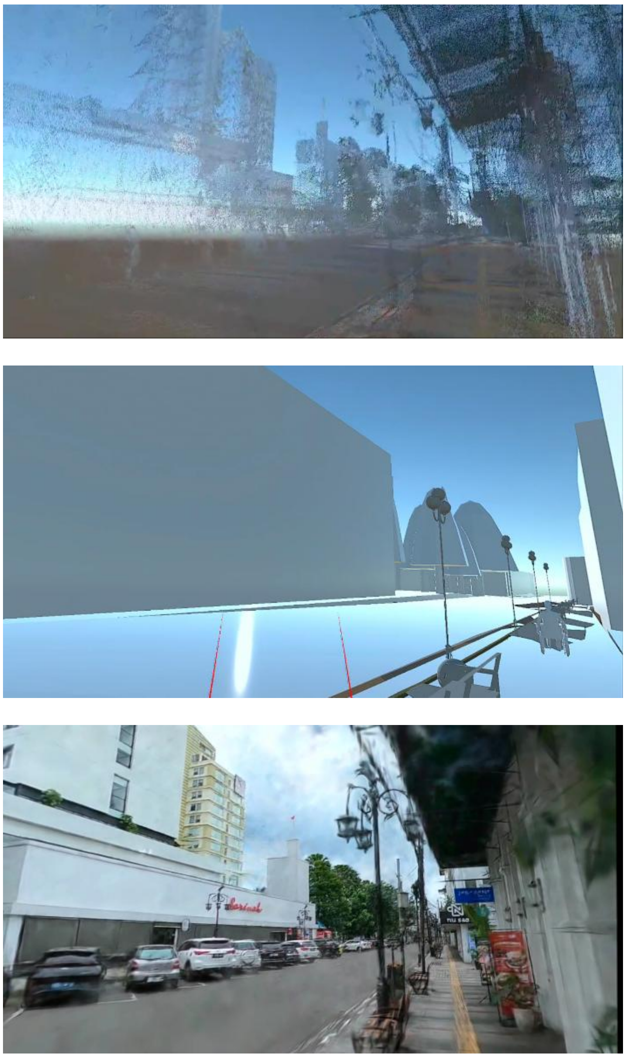

Figure 3. (a) LiDAR point cloud , (b) CityGML, and (c) 3DGS pedestrian model.

Beyond quantitative analysis, the integration of photorealistic data introduces an immersive and participatory dimension for evaluating pedestrian accessibility. The hybrid model was tested in an immersive assessment scenario involving wheelchair users. Figure 4 shows a participant exploring the reconstructed pedestrian environment using a VR/MR headset.

Within the immersive environment, users evaluate accessibility conditions from a human-scale perspective based on the defined indicators, including the presence of obstacles, surface continuity, and perceived slope. The first-person perspective allows users to identify potential barriers that may not be immediately evident in purely geometric representations. This demonstrates the added value of perception-driven assessment in complementing metric-based analysis.

Overall, the results indicate that the integration of LiDAR, 3DGS, and CityGML enables a comprehensive understanding of pedestrian infrastructure by combining metric accuracy, semantic structure, and visual perception within a single analytical framework.

Figure 4. Wheelchair user performing immersive accessibility assessment in a VR/MR environment.

3.1 Georeferencing and Co-registration Accuracy

This section evaluates the geometric consistency between the LiDAR point cloud and the image-based reconstruction, as well as the resulting 3DGS models. The analysis focuses on georeferencing accuracy and spatial agreement between metric and photorealistic representations, which are essential for reliable integration within the proposed multi-layer digital twin framework.

3.1.1 Georeferencing of Image-based Reconstruction to LiDAR

The first evaluation examines the georeferencing accuracy of the image-based reconstruction relative to the handheld LiDAR point cloud, which serves as the primary geometric reference due to its centimeter-level accuracy. The reconstruction, initially defined in a relative coordinate system, was transformed into the LiDAR coordinate system using a 3D similarity transformation based on control points derived from features identifiable in both datasets.

The georeferencing accuracy was evaluated using residual errors at the control points during bundle adjustment. The results indicate root mean square (RMS) residuals of approximately 22.38 cm in the horizontal (XY) direction and 14.93 cm in the vertical (Z) direction, with a total three-dimensional RMSE of approximately 26.91 cm. These results reflect the expected accuracy of image-based reconstruction using control point-based georeferencing in complex pedestrian environments.

Analysis of individual control points shows that most points exhibit residuals below 25 cm, while several points present larger deviations. These discrepancies are primarily associated with areas characterized by complex geometry, limited feature visibility, or less favorable distribution of control points, which affect the stability of image-based matching and optimization.

Overall, the achieved accuracy is considered sufficient for pedestrian-scale applications, where the primary objective is to support integrated visualization and accessibility assessment rather than high-precision engineering tasks. The results confirm that the proposed georeferencing strategy provides a reliable spatial foundation for subsequent multimodal integration.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-1117-2026

3.1.2 Geometric Consistency between LiDAR and 3DGS

Following the georeferencing stage, the geometric consistency between the handheld LiDAR point cloud and the 3DGS representations was evaluated. Two 3DGS datasets were analyzed based on different input sources: (i) a LiDAR-based 3DGS model generated from fisheye imagery associated with the handheld LiDAR system, and (ii) a GoPro-based 3DGS model generated from 360° video imagery.

To quantify geometric consistency, a cloud-to-cloud distance analysis was performed between the LiDAR point cloud and each 3DGS dataset. The results show that the LiDAR-based 3DGS model achieves a mean spatial deviation of approximately 0.38 m, with a standard deviation of 0.34 m. In contrast, the GoPro-based 3DGS model exhibits a higher mean deviation of approximately 0.57 m, with a comparable standard deviation of 0.35 m.

The distribution of distances, as illustrated in Figure 5a and 5b, indicates that the LiDAR-based 3DGS representation contains a higher proportion of points with smaller deviations, while the GoPro-based model shows a broader distribution with increased variability across the scene.

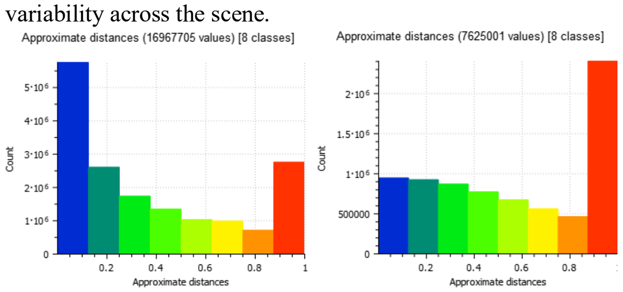

Figure 5. Cloud-to-cloud distance distributions between the LiDAR point cloud and (a) the LiDAR-based 3DGS model, and (b) the GoPro-based 3DGS model.

These differences are primarily attributed to the characteristics of the input data. The LiDAR-based 3DGS benefits from more stable image acquisition that is intrinsically linked to the LiDAR trajectory, resulting in improved geometric consistency. In contrast, the GoPro-based reconstruction relies on 360° video imagery, which is more sensitive to motion, lower positional accuracy, and variations in feature visibility, leading to increased geometric discrepancies.

Despite these differences, both 3DGS representations remain spatially aligned with the LiDAR reference and do not exhibit significant systematic distortions. This indicates that the proposed integration framework supports the coexistence of metric and photorealistic representations within a unified spatial reference system.

These findings highlight a clear trade-off between geometric accuracy and visual coverage. While LiDAR-based imagery provides higher spatial consistency, 360° video data offer richer contextual information and broader scene coverage. This trade-off is particularly relevant for pedestrian accessibility assessment, where both metric precision and perceptual understanding are required.

3.2 Pedestrian Accessibility Assessment

This section presents the evaluation of pedestrian sidewalk conditions using three complementary approaches, namely LiDAR-based analysis, field survey observations, and immersive assessment within a Virtual Reality (VR) environment. The integration of these approaches enables a comprehensive understanding of pedestrian accessibility by combining metric measurements, real-world observations, and human-scale visual inspection.

3.2.1 LiDAR-based Assessment

The LiDAR point cloud was used as the primary dataset for quantitative assessment of sidewalk conditions due to its high geometric accuracy. Geometric indicators such as surface continuity, elevation differences, slope, and the presence of obstacles were analyzed directly from the point cloud data.

The analysis enables the identification of physical characteristics affecting pedestrian accessibility, including uneven surfaces, abrupt height changes, and spatial obstructions caused by street furniture, bollards, and vegetation. The centimeter-level accuracy of the LiDAR data ensures reliable detection of small-scale geometric variations that may influence pedestrian movement and accessibility.

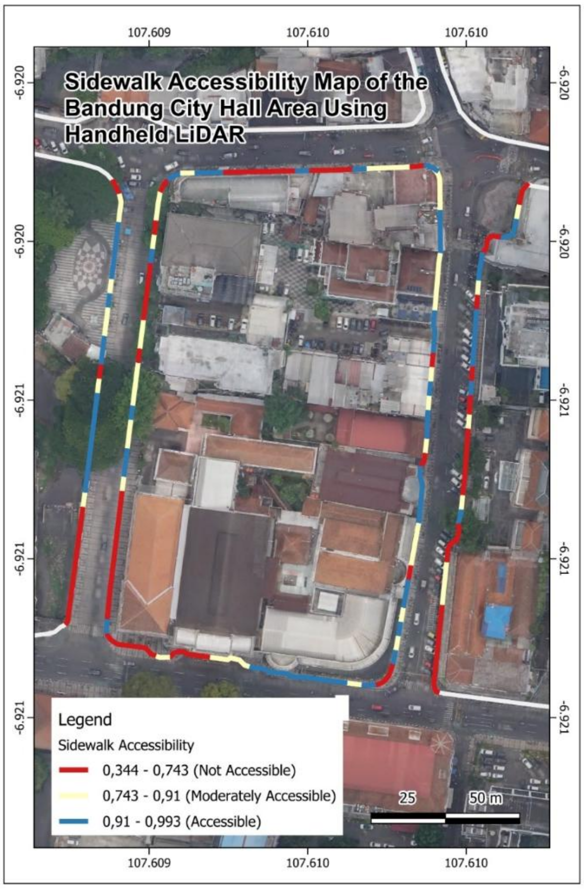

Figure 6. LiDAR-based pedestrian accessibility map derived from geometric indicators (slope, roughness, and obstacles).

The extracted indicators were integrated into an accessibility index, which was subsequently classified into three categories: accessible, moderately accessible, and not accessible, forming the basis for comparison with other assessment approaches. The resulting LiDAR-based accessibility map is shown in Figure 6.

3.2.2 Field Survey Assessment

Field-based observations were conducted along the same pedestrian corridor to document actual sidewalk conditions and accessibility barriers. The survey focused on identifying physical obstacles, surface conditions, and usability constraints experienced by pedestrians.

This approach captures real-world conditions that may not be fully represented in geometric data alone, such as temporary obstructions, surface degradation, and contextual factors affecting pedestrian movement. The field survey serves as a reference for validating the findings obtained from LiDAR-based and VR-based assessments.

The field observations were performed based on the same pedestrian segments used in the LiDAR analysis, ensuring consistency in spatial units for comparison. The accessibility map derived from field survey observations is presented in Figure 7.

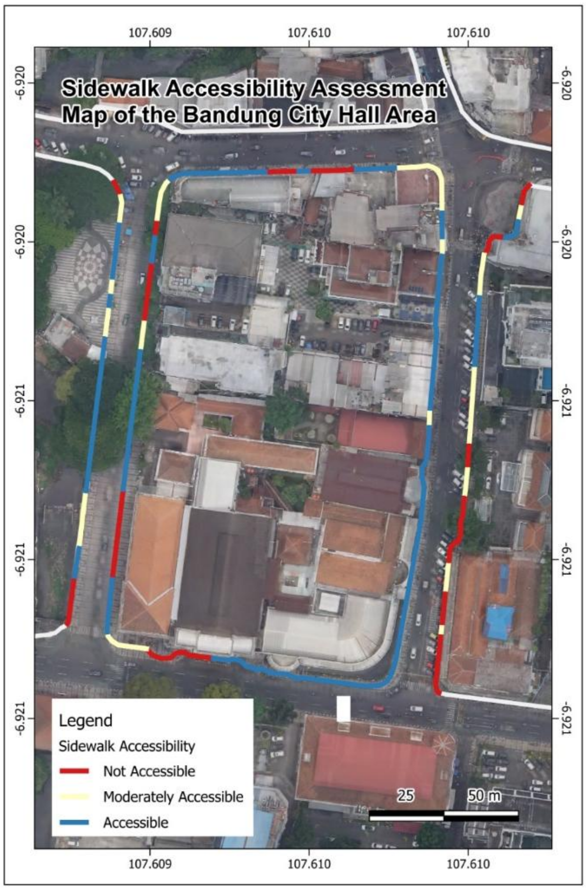

Figure 7. Field survey-based pedestrian accessibility map representing observed sidewalk conditions.

3.2.3 VR-based Assessment

An immersive assessment was conducted within the Virtual Reality (VR) environment using the integrated multi-layer datasets, including the LiDAR point cloud, 3DGS model, and CityGML representation. The VR environment enables users to explore the pedestrian space at a human scale and visually inspect sidewalk conditions.

Visual inspection in VR allows the identification of accessibility issues such as narrow pathways, obstructed routes, and surface irregularities. The photorealistic representation provided by 3DGS enhances the perception of environmental context, while the LiDAR data ensures spatial consistency.

Users performed the assessment based on visual interpretation of accessibility indicators, including perceived slope, surface continuity, and the presence of obstacles. This approach enables intuitive evaluation of pedestrian infrastructure conditions that are difficult to capture through purely geometric analysis. The accessibility map obtained from the VR-based assessment is shown in Figure 8.

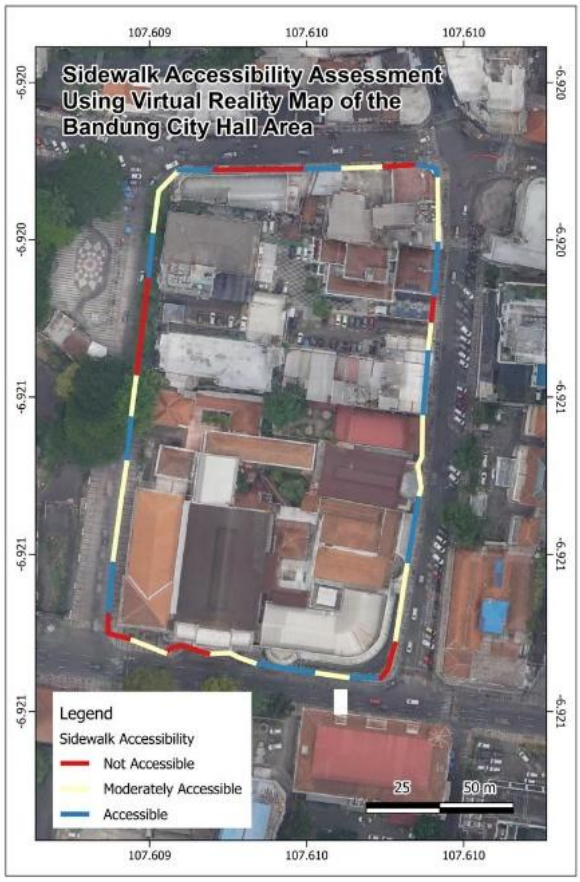

Figure 8. VR-based pedestrian accessibility map derived from immersive evaluation using 3DGS.

3.2.4 Comparative Analysis

As illustrated in Figures 6–8, the spatial distribution of accessibility levels shows both similarities and differences across the three assessment approaches.

The results obtained from the three assessment approaches were compared using the LiDAR-based classification as the reference, due to its higher geometric reliability. The comparison was performed at the pedestrian segment level using confusion matrices to quantify agreement between methods.

The field survey assessment achieves an overall accuracy of approximately 85.7%, indicating a strong agreement with the LiDAR-based results. This high consistency demonstrates that direct field observation aligns well with geometric measurements, with minor discrepancies primarily caused by subjective interpretation and temporal variations in field conditions.

In contrast, the VR-based assessment achieves an overall accuracy of approximately 75.4%, which is lower than the field survey results. The confusion matrix indicates that most discrepancies occur in distinguishing between moderately accessible and not accessible segments. These differences are primarily influenced by limitations in depth perception, occlusions, and the absence of precise metric cues in the immersive environment.

Despite this lower accuracy, the VR-based approach provides significant advantages in terms of accessibility, efficiency, and user engagement. The immersive environment allows users to evaluate pedestrian conditions remotely and from a human-scale perspective, which is particularly valuable for participatory and inclusive assessment scenarios.

Overall, the comparison highlights a complementary relationship between the three approaches. LiDAR-based analysis provides reliable quantitative measurement, field surveys offer contextual validation, and VR-based assessment enhances perception-driven evaluation. The integration of these approaches enables a more comprehensive and robust assessment of pedestrian accessibility within a digital twin framework.

4. Conclusion

This study presented a hybrid framework integrating handheld LiDAR, 3D Gaussian Splatting (3DGS), and CityGML semantic modelling for pedestrian-scale accessibility assessment within an urban digital twin environment. The proposed approach combines metric accuracy, photorealistic representation, and semantic structure to support both quantitative analysis and perception-driven evaluation of pedestrian infrastructure.

The results demonstrate that LiDAR-based analysis provides the most reliable representation of pedestrian accessibility due to its high geometric accuracy and ability to capture detailed surface characteristics. Field survey assessment shows strong agreement with LiDAR results, achieving an overall accuracy of approximately 85.7%, confirming the validity of the LiDAR-derived accessibility model.

In contrast, VR-based assessment achieves a lower agreement of approximately 75.4%, primarily due to limitations in depth perception, occlusions, and the absence of precise metric cues in immersive environments. However, VR offers significant advantages in terms of accessibility, efficiency, and user engagement, enabling remote and interactive evaluation of pedestrian conditions.

Importantly, the findings highlight the potential of VR-based assessment to support participatory accessibility evaluation. By enabling immersive interaction with the urban environment, VR allows users, including people with disabilities, to actively contribute to the assessment process without requiring physical presence in the field. This capability is particularly relevant for inclusive urban planning and citizen-centered digital twin applications.

Overall, the results indicate that LiDAR, field survey, and VR-based approaches are not interchangeable, but complementary. LiDAR provides reliable quantitative measurement, field surveys offer contextual validation, and VR enables perception-driven and participatory evaluation from a human-scale perspective.

The proposed multi-layer framework demonstrates the potential of integrating geometric, semantic, and photorealistic representations to support inclusive urban analysis. Future work will focus on improving semantic automation, enhancing metric reliability in radiance-field representations, and developing real-time digital twin systems that support continuous and participatory accessibility monitoring.

References

Beil, C., Ruhdorfer, R., Coduro, T., Kolbe, T. H., 2020. Detailed streetspace modelling for multiple applications: Discussions on the proposed CityGML 3.0 transportation model. ISPRS International Journal of Geo-Information, 9(10), 603.

Fadilah, W. A., Murtiyoso, A., Landes, T., Grussenmeyer, P., 2026. Metric assessment of 3D Gaussian splatting for UAV-based urban heritage reconstruction. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-2/W12-2026, 143–150.

Jawset Visual Computing, 2024. Postshot: AI-powered photogrammetry and neural rendering software. Available at: https://www.jawset.com/ (accessed 5 March 2026).

Jiang, C., Gao, R., Shao, K., Wang, Y., Xiong, R., Zhang, Y., 2025. LI-GS: Gaussian splatting with LiDAR incorporated for accurate large-scale reconstruction. IEEE Robotics and Automation Letters, 10(2), 1864–1871.

Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G., 2023. 3D Gaussian splatting for real-time radiance field rendering. ACM Transactions on Graphics (TOG), 42(4), Article 139.

Miura, S., Yoshida, M., Nakamura, F., Wongwiriya, P., Tanaka, S., Ariyoshi, R., 2024. Spatial and timeframe distribution of diverse pedestrian activities in the street network of Southeast Asian developing cities. Journal of Asian Architecture and Building Engineering, 23(1), 386–404.

Scolamiero, V., Boccardo, P., La Riccia, L., 2025. Mobile mapping system for urban infrastructure monitoring: Digital twin implementation in road asset management. Land, 14(3), 597.

Suwardhi, D., Agustina, L. K., Andika, W., Widyastuti, R., Fadilah, W. A., Murtiyoso, A., Remondino, F., 2026. Integrating terrestrial laser scanning and 3D Gaussian splatting for heritage building documentation and monitoring. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-2/W12-2026, 463–470.

Xiong, B., Zheng, N., Liu, J., Li, Z., 2024. GauU-Scene V2: Assessing the reliability of image-based metrics with expansive LiDAR image dataset using 3DGS and NeRF. arXiv preprint, arXiv:2404.04880.

Acknowledgements

This research is funded by the Indonesian Endowment Fund for Education (LPDP) on behalf of the Indonesian Ministry of Higher Education, Science and Technology and managed under the EQUITY Program (Contract No. 4298/B3/DT.03.08/2025). This work is also supported by the Ford Foundation through a disability-inclusive research initiative conducted in collaboration with Pergerakan Penyandang Disabilitas dan Lanjut Usia (DILANS) Indonesia.