Towards Smart Airports: Use of Segmentation and Classification Algorithms to Map Airdromes Features using High-Spatial Resolution LiDAR and Photogrammetry

Aluizio Brito Maia[1], Ezequiel Silva Rocha[1], Elton Vicente Escobar Silva[2], Rômulo Marques-Carvalho[3]
Cláudia Maria de Almeida[4], Eduardo Moraes Arraut[1], Evandro José da Silva[1]

[1]Aeronautics Institute of Technology, Brazil, {aluizio, ezequielesr, emarraut, evandro}@ita.br
[2]National Center for Early Warning of Natural Disasters, Brazil (Cemaden), eltescobar@gmail.com
[3]University of São Paulo (USP), Brazil, romulomarques@usp.br
[4]National Institute for Space Research (INPE), Brazil, claudia.almeida@inpe.br

Keywords: Smart Airports, Digital Twins, LiDAR, Photogrammetry, Hierarchical Segmentation, Random Forest.

Abstract

The digital replication of critical infrastructure is fundamental to the development of Smart Airports and the implementation of digital twins. Urban aerodromes, such as Congonhas Airport (SBSP) in São Paulo, face unique spatial constraints due to dense surrounding urbanization, requring accurate geometric models to manage operational risks and evaluate environmental impacts. Despite the potential of combining high-resolution photogrammetry and LiDAR, translating these massive datasets into structured 3D models presents significant methodological challenges. This study proposes an automated, multi-scale hierarchical segmentation and classification workflow to map aerodrome features using 0.25 cm spatial-resolution orthophotos and high-density LiDAR data. The methodology employs the Felzenszwalb-Huttenlocher algorithm across three spatial levels: macro-scale functional zones, elevated infrastructure via normalized Digital Surface Models (nDSM), and ground markings employing LiDAR intensity data. A Random Forest classifier was then applied to categorize thirteen distinct airport classes. The hierarchical approach outperformed traditional single-scale segmentation, achieving an F1-Score of 0.87 compared to the baseline of 0.53. The integration of LiDAR data proved crucial for distinguishing spectrally similar features, such as gray hangars on asphalt pavements. The proposed workflow automates the extraction of structural features, providing a foundational geometry for predictive maintenance, spatial organization, and urban conflict modeling.

1. Introduction

Analyses of critical urban infrastructures require highly realistic representations of their constituent components. For instance, when modelling the impact of natural disasters on a bridge or building, it is essential that urban models integrate dynamic virtual representations capable of precisely reflecting the physical and operational states of the infrastructure, thereby establishing a fundamental basis for the development of digital twins (Hamalainen, 2021).

Digital twins enable comprehensive analyses across multiple disciplines by assimilating specifications of a real-world object's diverse components into a cohesive workflow (Zhang, 2024). Consequently, remote sensing and digital twin technologies are inherently interconnected; developing a meticulous digital replica requires specialized sensors to capture spatial and material data. Photogrammetry and LiDAR have emerged as crucial technologies for this purpose, owing to their ability to capture the geometric structure and material properties of complex targets (Xue et al., 2020; Mohammadi et al., 2021; Rocha et al., 2026).

Among the spatial entities positioned to benefit from these advancements, airports stand out as prime candidates for digital replication. This transition enables the precise modeling of critical assets, including passenger terminals, hangars, and control towers. Furthermore, this replication serves as the methodological cornerstone for developing "Smart Airports", ecosystems in which engineering and management operations rely on continuous, precise spatial data (Conde et al., 2022).

In this study, a Smart Airport is defined as an aerodrome ecosystem in which physical infrastructure, operations, and management decisions are supported by continuously updated, geometrically accurate, and semantically classified spatial data, i.e., a digital twin of the airfield, enabling data-driven security planning, predictive maintenance, and risk management.

A viable approach to generating these representations involves integrating airborne laser scanning (LiDAR) and high-resolution photogrammetry, processed with segmentation and classification algorithms. While LiDAR provides an accurate three-dimensional geometric framework, photogrammetry provides spectral and textural information that is crucial for material identification. Automating this workflow to achieve sufficient detail enables airport managers to conduct security planning and predictive maintenance with greater confidence than traditional methods permit (Xue et al., 2020; Mohammadi et al., 2021).

Despite the potential of these technologies, converting extensive datasets derived from LiDAR and photogrammetry into structured digital twins presents various methodological challenges. The creation of 3D building models requires determining an appropriate level of detail that balances computational efficiency with the required architectural accuracy. Additionally, integrating these remote sensing outputs into comprehensive frameworks, such as Building Information Modeling (BIM), is fundamental to endowing raw geometric data with semantic meaning, thereby transforming point clouds into intelligent infrastructure assets (Rashdi et al., 2022).

A significant obstacle within this workflow is the transition from raw data acquisition to fully modeled entities. The manual extraction of structural features is labor-intensive and susceptible to human error, highlighting the necessity for automated processing pipelines. To address the geometric complexity of airport infrastructure, ranging from runways to specific terminal rooftop, Object-Based Image Analysis (OBIA) and hierarchical segmentation methods have proven to be considerably more effective than conventional pixel-based techniques (Men et al., 2020). By clustering pixels and point cloud data into meaningful spatial objects, certain algorithms can automatically identify, separate, and delineate specific architectural elements, such as roofs, facades, and runway pavements. This object-oriented methodology facilitates the 3D reconstruction of airport facilities (Han et al., 2020).

Nevertheless, a considerable gap remains in methodologies for the automated 3D modeling of aeronautical infrastructure by integrating LiDAR and photogrammetric datasets. Although general urban modeling techniques are available, airports exhibit distinctive geometric and regulatory constraints that require specialized solutions. Addressing this gap is crucial for enhancing the operational readiness of airport digital twins (Conde et al., 2022; Lv et al., 2007; Brassel et al., 2020). Therefore, the main aim of this research is to propose and assess an automated, multi-scale workflow for the segmentation and classification of aeronautical features using high-resolution orthophotos and LiDAR data. The research problem here is that manually extracting and classifying aerodrome infrastructure from high-resolution photogrammetric and LiDAR data is labour-intensive, error-prone, and does not scale to the frequent updates a digital twin requires. in this case, our study is applied in the Congonhas Airport (SBSP), an urban aerodrome with dense surrounding development, is used as a representative case of the geometric and spectral complexity found in constrained urban airfields. So, for this, the research question of this study is: Does a multi-scale hierarchical segmentation strategy outperform a conventional single-scale segmentation for delineating aerodrome features at both macro (functional zones) and micro (surface markings) scales? and: Can a Random Forest classifier, trained on segmentation objects derived from orthophoto, nDSM, and LiDAR intensity data, reliably discriminate spectrally similar aerodrome classes (e.g., hangars vs. asphalt pavements)?

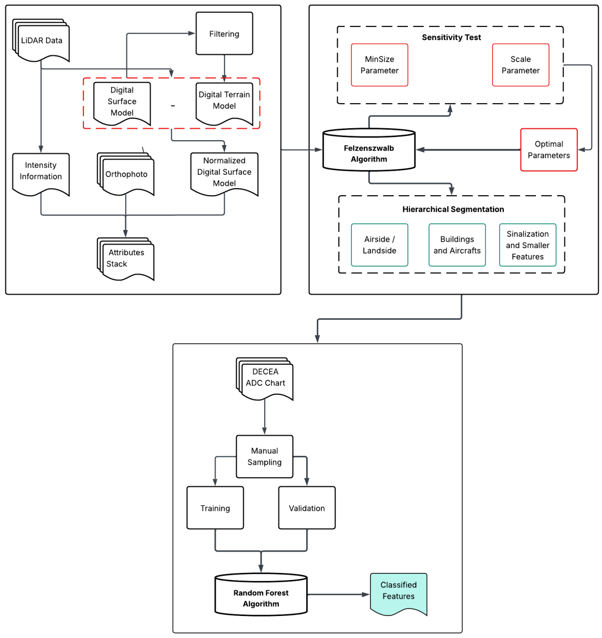

Figure 1. General Methodology.

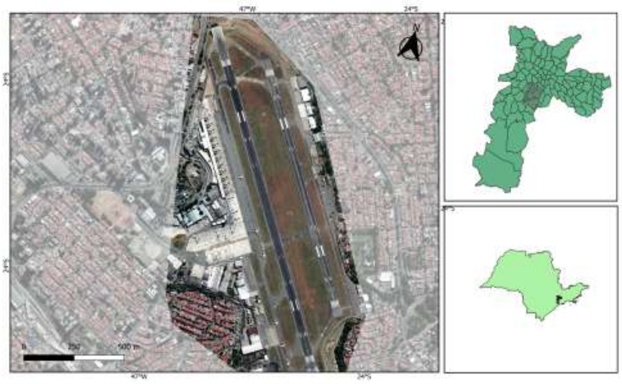

Figure 2. Congonhas Airport (SBSP), located in São Paulo, Brazil.

2. Material and Methods

The general methodology flowchart is presented in Figure 1.

2.1 Study Area

Congonhas Airport (SBSP), located in São Paulo (Figure 2), is the second-busiest airport in Brazil and serves as a vital domestic hub, emphasizing its strategic significance within the national aviation sector. Established in the 1930s, it maintained the status of the country's busiest airport for 16 consecutive years, from 1990 to 2006 (Dias, 2020).

Initially built in a sparsely populated region, the swift expansion of urban development over the decades has completely encroached upon the facility. This has resulted in a centrally located airport, nowsurrounded by dense urban growth and skyscrapers. Sucha a distinctive spatial configuration makes SBSP an exemplary case study, demonstrating how essneital infrastructure can significantly impact and be limited by the evolving urban landscape (Scatolini and Alves, 2016). Given these dense urban surroundings, spatial organization and operational planning must be carefully managed to mitigate risks and prevent accidents. Consequently, this aerodrome was selected as the primary study area for this research.

2.2 Remote Sensing Data and Pre-Processing Steps

The dataset consists of RGB (Red, Green, and Blue) image with a spatial resolution of 0.25 cm and a total size of 11 GB. The LiDAR data, characterized by a density of 19 points per square meter, has a file size of 630 MB.

In the pre-processing stage, a Digital Surface Model (DSM) was generated from a point cloud comprising 128,945,840 points. The derived raster has a spatial resolution of 0.25, calculated by averaging the intensity of points within each cell to ensure consistency with the ortophoto. This procedure required approximately 50 seconds on a workstation equipped with an Intel Core i9-13900HX (2.20 GHz), 32 GB of RAM, and an NVIDIA GeForce RTX 4070 GPU, operating on Windows 11.

Subsequently, a Digital Terrain Model (DTM) was derived from the DSM by filtering out building footprints using OpenStreetMap (OSM) data OpenStreetMap contributors (2026) A normalized Digital Surface Model (nDSM) was then computed by subtracting the DTM from the DSM ($nDSM = DSM - DTM$). This process resulted in a raster that represents the absolute heights of features, effectively eliminating terrain effects.

Then, as input attributes for the segmentation, a stack comprising the ortophoto bands, the intensity data, and the nDSM was utilized.

2.3 Hierarchical Segmentation and Classification Process

Image segmentation was performed using the Felzenszwalb-Huttenlocher algorithm (Felzenszwalb and Huttenlocher, 2004), a graph-based methodology that partitions the image by comparing variations between neighboring pixels. This technique is notably efficient for extensive datasets, as it preserves significant image details within low-contrast areas while exhibiting minimal computational demand. The parameters of the algorithm were meticulously calibrated to ensure that the derived segments remained consistent with the 0.25 cm spatial resolution of the orthophoto.

The algorithm was implemented in Python using the scikit-image library. To identily the optimal configuration, a sensitivity analysis was performed through a grid search approach. Systematic evaluation of variations in Scale, Sigma (σ), and Min_{Size} within ranges compatible with the 0.25 cm spatial resolution was performed. This process ensured that the chosen parameters achieved the optimal equilibrium between preserving fine-scale details and preventing over-segmentation.

An initial assessment of the optimal segmentation parameters was conducted by manually sampling 35 critical aerodrome features. This evaluation used the Aerodrome Chart (ADC) (see Figure 3), provided by the Brazilian Department of Airspace Control (DECEA), as a reference (Department of Airspace Control DECEA, 2026). The 35 reference features used to assess segmentation quality were selected to cover all functional zones of the aerodrome (airside and landside) and all spatial scales addressed by the hierarchical approach (macro functional zones, elevated infrastructure, and surface markings), rather than being confined to a single sector of the airport. The DECEA Aerodrome Chart (ADC) follows ICAO Annex 4 (Aeronautical Charts) accuracy and resolution requirements for aerodrome obstacle and infrastructure data; however, DECEA does not publish a quantitative positional accuracy value for the published chart itself. The ADC was therefore used as the best-available official reference for aerodrome feature identity and nominal location, rather than as a surveyed ground-truth dataset with a known error budget. This distinction is acknowledged as a limitation of the validation strategy adopted in this study.

To evaluate segmentation, the Jaccard similarity index was employed to quantify the spatial overlap between two datasets. The index ranges from 0 (no overlap) to 1 (complete overlap). The Jaccard index is given by:

$$J = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

where:$|A \cap B|$: Number of elements shared between A and B (intersection).$|A \cup B|$: Total number of unique elements in A and B (union).

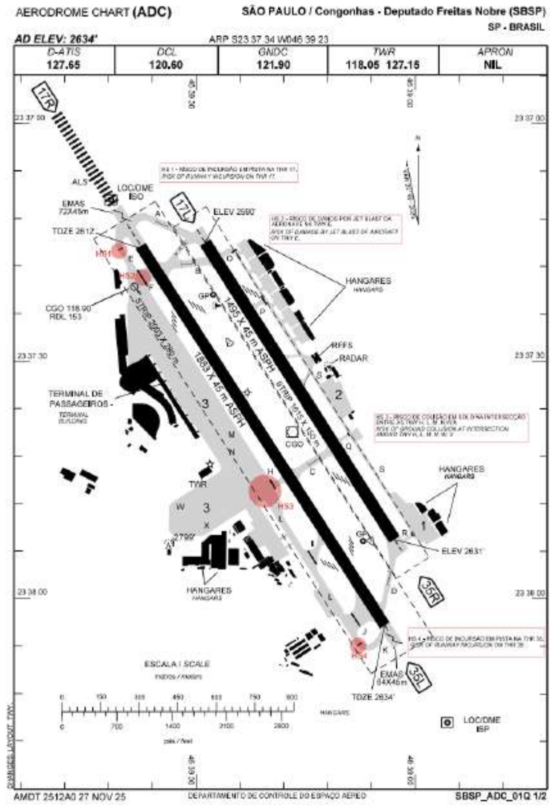

Figure 3. The Aerodrome Chart (ADC) accompanying the Aeronautical Information Publication of Congonhas Airport (SBSP).

Once the optimal parameters were determined, a multiscale hierarchical segmentation methodology was implemented (Figure 4). This approach organizes the segmentation process into three primary levels:

- Level 1 (Macro-scale): Parameterization designed to segregate extensive functional zones, specifically distinguishing the airside from the landside.
- Level 2 (Buildings): Integration of the DSM to for the extraction of elevated infrastructure, including hangars and passenger terminals.
- Level 3 (Markings and Smaller Structures): Incorporation of LiDAR intensity data to precisely delineate runway and taxiway surface markings and other structures.

This hierarchical approach designed to effectively capture the diverse spatial dimensions of aerodrome features, thereby preventing the over-segmentation of extensive homogeneous areas while preserving the fine details of smaller targets. Finally, its performance was evaluated in comparison to a conventional single-scale segmentation baseline to assess the enhancements in feature extraction precision.

The classification process was conducted using the Random Forest (RF) algorithm (Breiman, 2001), implemented in the R programming environment via the randomForest package (Liaw and Wiener, 2002). In order to enhance the model's effectiveness, a sensitivity analysis was performed to fine-tune the hyperparameters *mtry* and *ntree*. The optimal configuration was identified by determining the lowest Out-of-Bag (OOB) error,

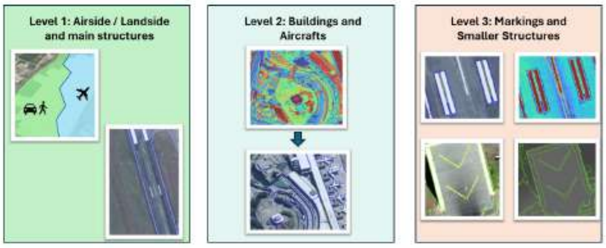

Figure 4. Hierarchical segmentation approach.

an inherent metric offering an unbiased estimate of the model's internal classification accuracy. The algorithm was trained on a dataset consisting of 615 samples, distributed among 13 classes, which were selected referencing the DECEA Aerodrome Chart (ADC) as the ground truth.

3. Results and Discussion

The comparison of segmentation outcomes following the parameter sensitivity analysis is illustrated in Figure 5. The assessment of the single-scale segmentation indicated a tendency towards over-segmentation. This occurrence is common in Geographic Object-Based Image Analysis - GEOBIA workflows, wherein the initial segmentation is designed to ensure that no target is overlooked, delegating the merging and refinement of objects to the classification and post-processing stages. The parameter configuration of $scale = 1000$, $\sigma = 0.5$, and $min\_size = 2000$ was selected, as it provided the optimal balance between edge preservation and the reduction of excessive fragmentation. Conversely, Figure 5(B) displays the results of the hierarchical approach, exhibiting more cohesive delineation of shapes and a substantial reduction in over-segmentation.

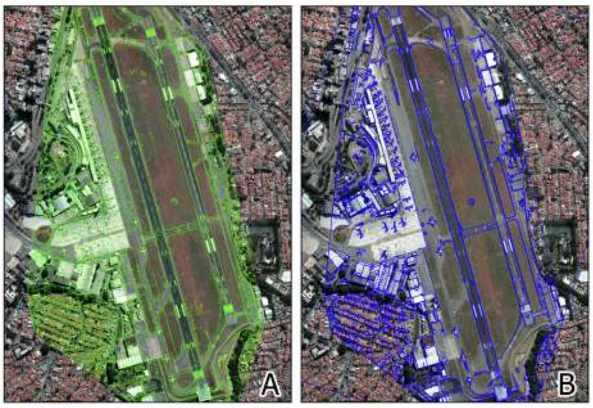

Figure 5. Comparison between segmentation approaches with the optimal parameters. (A) indicates a traditional single-scale segmentation and (B) indicates an hyerarquical approach.

Following the segmentation phase, the classification algorithm was applied, and the comparative results are illustrated in Figure 6. It is evident that the traditional single-scale method resulted in higher misclassification rates for larger features (A); for instance, the runway was frequently confused with the Engineered Materials Arresting System (EMAS) and taxiways. While the overall classification performance improved substantially, increasing the macro F1-score from 0.53 to 0.87, precise boundary delineation remained a challenge, with several taxiway edges being misclassified as runways. Furthermore, there was increased confusion between other airport facilities and the main terminal, primarily due to their similar spectral signatures in the orthophoto and comparable elevation profiles within the DSM.

The figure 7 shows the extraction results for individual aerodrome features using the hierarchical classification model. The visual evaluation corroborates the model's capacity in delineating targets across multiple scales. Macro-structures, such as the main runway (A) and the runway strip (J), were effectively isolated. Moreover, the incorporation of LiDAR intensity data demonstrated significant potential for micro-scale features, facilitating the accurate extraction of critical surface markings, including the threshold (H) and aiming points (I).

Nevertheless, the visual results additionally emphasize particular challenges inherent to aerodrome environments. Features that sharecontinuous paved surfaces and posses similar spectral signatures (the taxiways (B) and apron areas (E)) exhibited higher levels of edge fragmentation. Additionally, although structural classification was predominately successful (e.g., the EMAS in G and hangars in D), the precise delineation of boundaries between adjacent infrastructure elements continues to be an area for subsequent post-processing refinement.

4. Conclusion

This study presents a segmentation-based approach for the extraction of aerodrome infrastructure. The results indicate that the proposed hierarchical segmentation technique effectively delineates delineate aerodrome features, demonstrating computational, geometric, and operational viability for the development of airport digital twins. Furthermore, this research advocates for the implementation of advanced classification models tailored to air transport management. By integrating multi-source data within the expansive framework of geospatial technologies and smart cities, it provides a comprehensive product with diverse informational value.

The workflow proposed here directly supports three characteristics commonly required of a digital twin of critical infrastructure: (i) geometric fidelity, through the sub-decimetre orthophoto and high-density LiDAR point cloud (19 pts/m^2) used as the geometric backbone; (ii) semantic richness, through the 13-class object-based classification, which converts undifferentiated point-cloud/raster data into named, queryable infrastructure entities (runway, EMAS, hangars, terminal, etc.); and (iii) updatability, since the automated segmentation-classification pipeline can, in principle, be re-run on new epochs of orthophoto/LiDAR acquisition to refresh the model without manual re-digitizing.

Although the results are promising, some limitations need to be recognized. A more thorough comparison of different parameter settings is necessary to fully grasp their effects. Moreover, the dependence on time-consuming training sample collection methods continues to be a challenge that must be tackled.

Future research should compare different segmentation algorithms with varying computational demands to assess their effectiveness. Incorporating advanced classifiers, especially Neural Networks and Deep Learning models, is also strongly advised to enhance classification accuracy. Additionally, upcoming stages will focus on visualizing the classification results

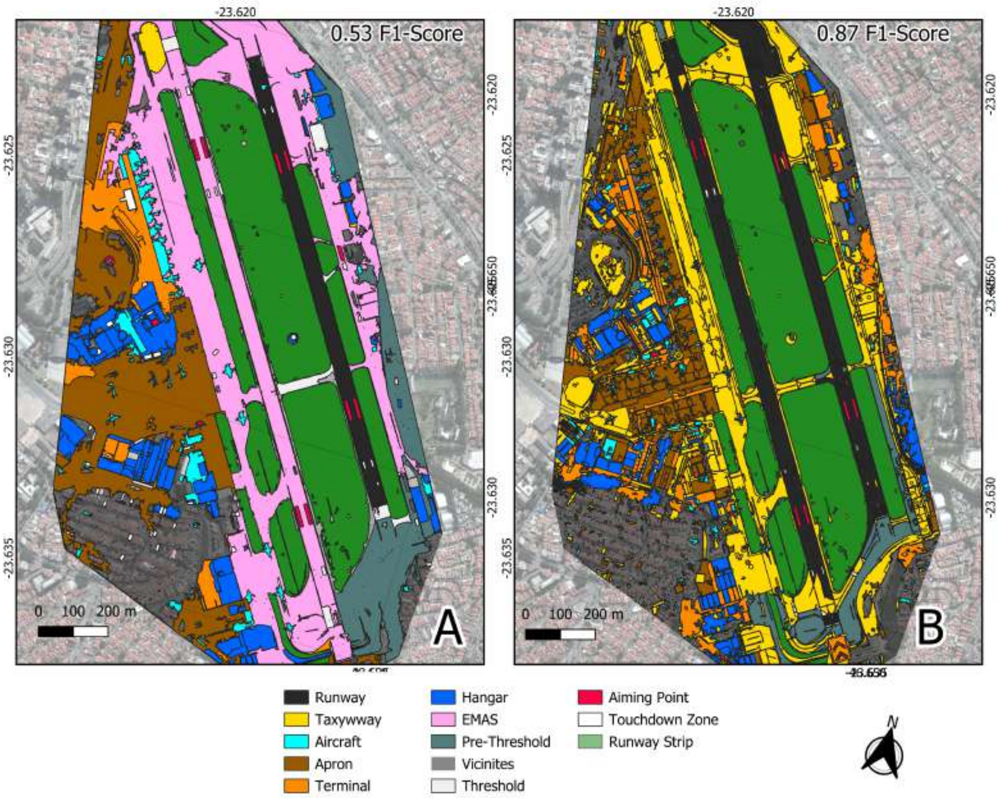

Figure 6. Visual comparison of classification outcomes using the optimal segmentation parameters. (A) Classification based on traditional single-scale segmentation (macro F1-score = 0.53), showing frequent confusion between the runway, the EMAS, and taxiways. (B) Classification based on the proposed three-level hierarchical segmentation (macro F1-score = 0.87), showing markedly reduced misclassification for these same classes

within a 3D environment using high-density LiDAR data. This visualization will be essential for a thorough understanding of the aerodrome's geometric layout and operational limitations.

A formal, quantitative uncertainty assessment (e.g., propagation of DSM/DTM vertical error into the nDSM, or confidence intervals on the reported F1-scores) was not conducted in this study and is acknowledged as a limitation. The only quality-control step applied was the visual cross-check between classified segments and the DECEA Aerodrome Chart (ADC) used as ground truth (Section 2.3). Future work should incorporate a structured accuracy-assessment protocol to quantify both thematic and positional uncertainty

Acknowledgments

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)

References

Brassel, H., Zouhar, A., Fricke, H., 2020. 3d modeling of the airport environment for fast and accurate lidar semantic segmentation of apron operations. *2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC)*, 1–10.

Breiman, L., 2001. Random Forests. *Machine Learning*, 45, 5–32.

Conde, J., Munoz-Arcentales, A., Romero, M., Rojo, J., Salvachúa, J., Huecas, G., Álvaro Alonso, 2022. Applying digital twins for the management of information in turnaround event operations in commercial airports. *Advanced Engineering Informatics*, 54, 101723. https://www.sciencedirect.com/science/article/pii/S1474034622001811.

Department of Airspace Control DECEA, 2026. Aerodrome Chart (ADC) - São Paulo / Congonhas, SP (SBSP). https://aisweb.decea.mil.br/?i=cartas. Accessed: Apr. 10, 2026.

Dias, M., 2020. Air Transportation in Brazil: São Paulo Congonhas Airport. *Global Scientific Journals*, 8(2), 3244–3252.

Felzenszwalb, P. F., Huttenlocher, D. P., 2004. Efficient Graph-Based Image Segmentation. *International Journal of Computer Vision*, 59(2), 167-181. https://doi.org/10.1023/B:VISI.0000022288.19776.77.

Hamalainen, M., 2021. Urban development with dynamic digital twins in Helsinki city. *IET Smart Cities*, 3(4), 201-210. https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/smc2.12015.

Han, Y., Ma, S., Xu, Y., He, L., Li, S., Zhu, M., 2020. Effective Complex Airport Object Detection in Remote Sensing Images Based on Improved End-to-End Convolutional Neural Network. *IEEE Access*, 8, 172652-172663.

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. *R News*, 2(3), 18–22.

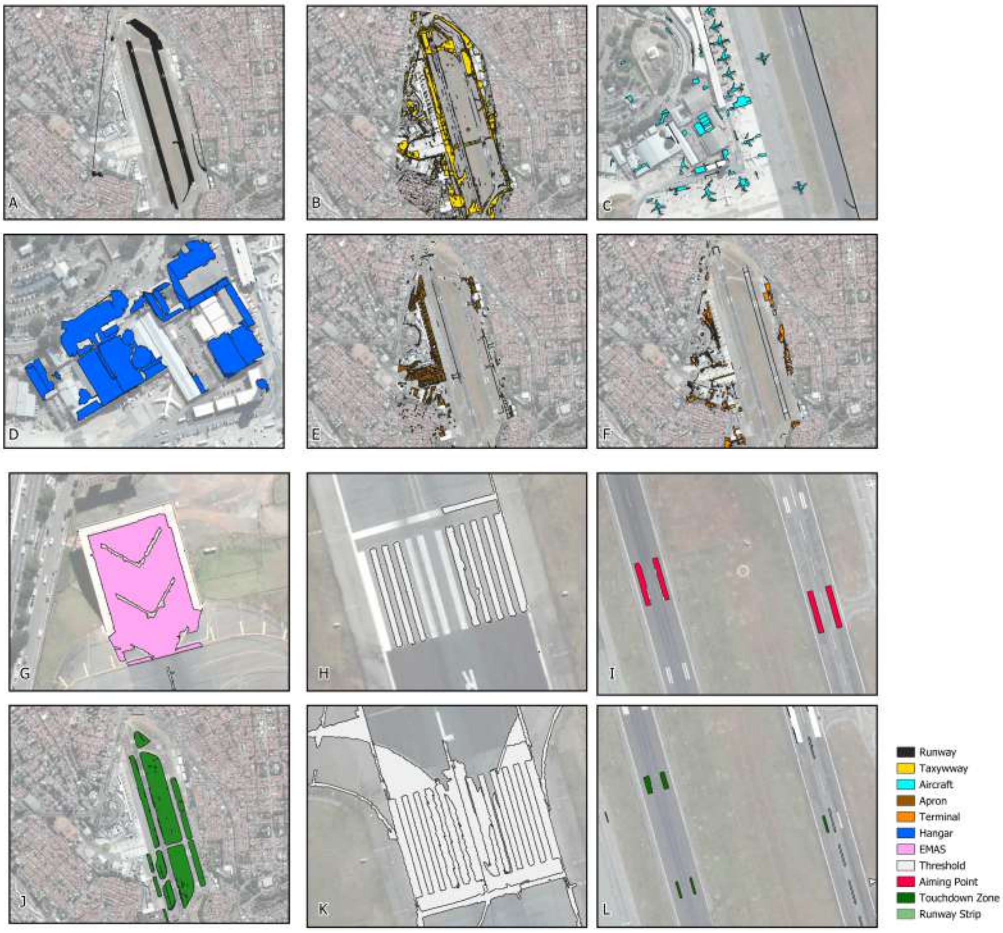

Figure 7. Results of the individual features.

Lv, Z., Sui, H., Zhang, X., Huang, X., 2007. Automatic updating and 3D modeling of airport information from high resolution images using GIS and LIDAR data. Y. Wang, B. Lei, J.-Y. Yang, J. Li, C. Wang, L.-P. Zhang (eds), *MIPPR 2007: Remote Sensing and GIS Data Processing and Applications; and Innovative Multispectral Technology and Applications*, 6790, International Society for Optics and Photonics, SPIE, 679018.

Men, Z. C., Jiang, J., Guo, X., Chen, L. J., Liu, D. S., 2020. AIRPORT RUNWAY SEMANTIC SEGMENTATION BASED ON DCNN IN HIGH SPATIAL RESOLUTION REMOTE SENSING IMAGES. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-3/W10, 361–366. https://isprs-archives.copernicus.org/articles/XLII-3-W10/361/2020/.

Mohammadi, M., Rashidi, M., Mousavi, V., Karami, A., Yu, Y., Samali, B., 2021. Quality Evaluation of Digital Twins Generated Based on UAV Photogrammetry and TLS: Bridge Case Study. *Remote Sensing*, 13(17). https://www.mdpi.com/2072-4292/13/17/3499.

OpenStreetMap contributors, 2026. OpenStreetMap database. https://www.openstreetmap.org. Acessado em: 23 ago. 2026.

Rashdi, R., Martínez-Sánchez, J., Arias, P., Qiu, Z., 2022. Scanning Technologies to Building Information Modelling: A Review. *Infrastructures*, 7(4). https://www.mdpi.com/2412-3811/7/4/49.

Rocha, E. S., Maia, A. B., de Almeida, C. M., Ruiz, P. R. D. S., 2026. Urban Morphology and Infrastructure Patterns: A LiDAR-Based 3D Cluster Analysis Using Verticalization as a Proxy. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-3/W4-2025, 313–318. https://isprs-annals.copernicus.org/articles/X-3-W4-2025/313/2026/.

Scatolini, F., Alves, C. J. P., 2016. Background noise analysis in urban airport surroundings of Brazilian cities, Congonhas Airport, São Paulo. *Revista de Saúde Pública*, 50. https://rsp.fsp.usp.br/article/background-noise-analysis-in-urban-airport-surroundings-of-brazilian-cities-congonhas-airport-sao-paulo/.

Xue, F., Lu, W., Chen, Z., Webster, C. J., 2020. From LiDAR point cloud towards digital twin city: Clustering city objects based on Gestalt principles. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, 418-431. https://www.sciencedirect.com/science/article/pii/S0924271620302070.

Zhang, Y., 2024. *Digital Twin: Architectures, Networks, and Applications*. Simula SpringerBriefs on Computing, 1 edn, Springer Cham.