Development of a High-resolution Geospatial Dataset for Sidewalk and Flood Analysis using Computer Vision

Andreza Lukosiunas, Lídia Raiza Sousa Lima Chaves Trindade, and Roberto Marcondes Cesar Junior
Institute of Mathematics and Statistics
University of São Paulo
São Paulo, SP, Brazil
Email: {andreza.lukosiunas, lidiaraiza, rmcesar}@usp.br

***Abstract*—Sidewalk infrastructure plays a vital role in advancing sustainable cities by enhancing mobility, promoting social inclusion, and strengthening urban resilience, particularly in the face of climate-related risks such as flooding. In several developing countries, sidewalks are not only of poor quality but also lack consistent and context-sensitive data that would allow for an adequate assessment of their condition. This paper presents the development of a high-resolution multimodal geospatial dataset aimed at analyzing sidewalk infrastructure in urban environments, with a focus on the borough of São Miguel in São Paulo, Brazil, an area marked by frequent flooding and socio-spatial inequalities. By integrating remote sensing, deep learning, and computer vision techniques, the study supports the adaptation and validation of semantic segmentation tools such as Tile2Net for the Brazilian context. The combination of these methods for sidewalk segmentation shows promising results, with preliminary findings demonstrating the effectiveness of Tile2Net integrated with dilation operation. The proposed approach contributes to scalable and socially equitable solutions for urban infrastructure assessment and inclusive territorial planning, with a particular focus on sidewalk networks in flood-prone areas, making them especially relevant for analyzing potential escape routes during such events.**

I. Introduction

Sidewalk infrastructure in large urban centers has emerged as a critical dimension for sustainable mobility, pedestrian safety, and social inclusion. Recent studies indicate that poor-quality or missing sidewalks disproportionately affect vulnerable groups, including people with disabilities, the elderly, and residents of peripheral areas [1]–[3]. In addition to hindering accessibility, these deficiencies exacerbate risks associated with extreme hydrological events such as floods and urban inundations, particularly in areas marked by socio-spatial inequalities [4], [5].

Given this context of vulnerability, advances in computer vision techniques, combined with remote sensing and deep learning, have expanded the possibilities for automated diagnosis of urban infrastructure. Recent methods enable the detection, classification, and semantic segmentation of urban elements with high accuracy from aerial or satellite imagery [6], [7]. Tools such as Tile2Net have gained prominence by employing multiscale hierarchical attention to semi-automatically segment and structure urban sidewalk networks from satellite imagery [8].

In this regard, the present study proposes the development of a high-resolution multimodal geospatial dataset focused on the analysis of sidewalk infrastructure in urban environments. The goal is to support and validate segmentation tools such as Tile2Net by adapting these methodologies to the Brazilian urban context, ultimately fostering more inclusive technological solutions aligned with the real needs of national cities. The initial focus is on the borough of São Miguel in São Paulo, an area characterized by a high incidence of floods and insufficient urban infrastructure.

II. Methodology

A. Region of Interest

The study will be conducted in the municipality of São Paulo, SP, Brazil. The capital is highly vulnerable to extreme hydrological events, including floods, flash floods, and other risks related to urban water dynamics [9]. According to data from GeoSampa [10], a total of 238 flood events and 351 flash flood occurrences have been recorded throughout the city to date (Figure 1). These events are spatially distributed across multiple districts, highlighting socio-spatial vulnerability and the lack of adequate stormwater drainage infrastructure and risk management measures.

For the purposes of dataset design and the initial application of segmentation methods, the borough of São Miguel was selected as the pilot area for this study. This choice is based on the high incidence of flooding and the reported deficiencies in urban infrastructure, particularly regarding sidewalks. Located in the eastern zone of the city, São Miguel encompasses districts such as Jardim Helena, Vila Jacuí and São Miguel [11], and presents a diverse socio-spatial profile, making it a representative case for assessing sidewalk conditions in contexts of social and environmental vulnerability.

The region also features a complex hydrological system, intersected by rivers and streams such as the Tietê River, and the Lajeado and Itaim streams. These characteristics, combined with high population density and infrastructural vulnerability, make São Miguel an ideal pilot area for testing sidewalk

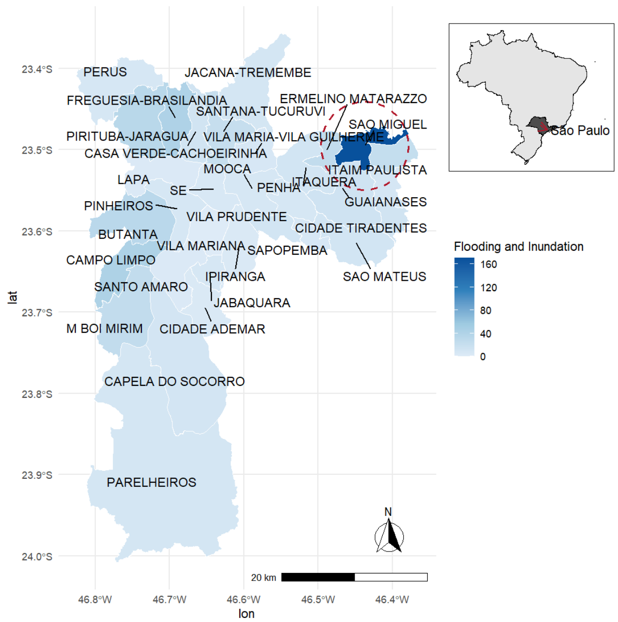

Fig. 1. Spatial distribution of flooding and flash flood occurrences in the city of São Paulo in 2025. Source: [10].

segmentation techniques and evaluating urban walkability in flood-prone zones.

B. Data Acquisition and Processing

A high-resolution multimodal geospatial dataset is being developed through the integration of multiple data sources with varying spatial resolutions and thematic contents, aiming to provide a robust foundation for the semantic segmentation of urban infrastructure. The dataset is organized hierarchically by borough, district, and image tiles, and incorporates both raster and vector data.

Among the raster data used, high-resolution RGB orthophotos (∼0.10 m/pixel) from the GeoSampa platform (reference year: 2020) stand out, offering sufficient visual detail for the accurate identification of sidewalks and related urban elements [10]. The orthophotos and associated shapefiles are originally projected using Brazil's official coordinate system, EPSG:31983 (SIRGAS 2000 / UTM zone 23S). To enable processing within the Tile2Net framework, the data were reprojected to EPSG:4326, the World Geodetic System (WGS 84). This global reference system is widely used in GPS technologies and is a standard requirement for many web-based mapping applications [12].

Additionally, multispectral images from the Sentinel-2 satellite, with a spatial resolution of 10 meters, were integrated into the analysis. These images are part of the Copernicus program of the European Space Agency (ESA) [13]. The spectral bands used included those required for the calculation of the Normalized Difference Water Index (NDWI), with acquisition dates corresponding to February 19, 2025 (rainy day) and July 19, 2025 (dry day). The images are provided at Level-2A processing, already atmospherically corrected, and were subsequently clipped to the geographical boundary of São Miguel. Table I presents the technical specifications of the imagery datasets used in this study.

TABLE I
TECHNICAL CHARACTERISTICS OF THE IMAGERY USED

Source	GeoSampa	Sentinel-2
Type	RGB	Multispectral
Year	2020	2022–2025
Res.	0.10 m	10 m, 20 m, 60 m
Bands	3 (RGB)	13
Format	JP2	TIFF/PNG
Freq.	One-time	5 days

An adapted version of the NDWI was employed, calculated using Sentinel-2 bands B03 (green) and B08 (near-infrared), as originally proposed by Gao [14]. Although initially developed to estimate vegetation moisture content, this index has been widely applied in urban areas to detect surface moisture and potential flooding, particularly in regions with sparse vegetation cover. The choice is justified by the 10 m spatial resolution available in both bands, which facilitates the detection of water accumulation on impervious surfaces. After calculation, the index was subjected to a *binarization process*, in which pixels with values greater than 0.00 were classified as *water*, while the remaining pixels were classified as *non-water*, resulting in the generation of seasonal flood masks. The total number of pixels classified as water was automatically computed in the R environment using the `global()` function from the `terra` package, which allows the extraction of global raster statistics. Subsequently, analyses using Synthetic Aperture Radar (SAR) data will be tested, enabling the monitoring of water extent under any weather or lighting conditions.

The dataset also incorporates socioeconomic and demographic data provided by the Brazilian Institute of Geography and Statistics (IBGE) and municipal databases, including information on population density, income distribution, education levels, and composite indices of social vulnerability. Vector layers complement the dataset, representing road networks, official administrative boundaries, land use and land cover classifications, and the location of public facilities (e.g., schools, health units, parks), as well as historically flood-prone zones.

Each georeferenced tile is documented with specific metadata, including central coordinates, area, pixel spatial resolution, and relevant contextual variables such as vegetation cover, slope, elevation, and proximity to essential services. For the purpose of training and validating segmentation models, the dataset also includes manually annotated ground truth data, enabling systematic evaluation of algorithmic performance.

C. Semantic Segmentation

Semantic segmentation of sidewalks was performed using the Tile2Net framework, with an additional dilation operation applied as a post-processing step to improve segmentation performance.

1) Tile2Net: the methodology adopted in this study is based on Tile2Net, a framework designed for sidewalk segmentation using orthorectified aerial imagery. The input data are organized into square tiles of 256×256 pixels, structured according to zoom level (z) and geographic coordinates—latitude (y) and longitude (x). Tile2Net generates three primary outputs related to sidewalk, road and crosswalk infrastructures:

- Semantic segmentation masks;
- Polygon representations;
- A topological sidewalk network.

The segmentation process is powered by a hierarchical multi-scale attention model [15], while polygonization is handled by the Python package Shapely. From these polygons, a sidewalk network is generated and will be used for further spatial analysis.

This methodology builds upon the original Tile2Net pipeline introduced in the seminal work Mapping the Walk [8]. Tile2Net has since been adopted in multiple recent studies on sidewalk detection and pedestrian network modeling [16]–[20].

2) Dilation: as Tile2Net has not yet been fine-tuned for the urban fabric of São Paulo, the segmentation results were enhanced by using a dilation operation applied to the road mask. This approach is based on the assumption that sidewalks emerge from the edges of roads. Incorporating prior information into the problem [21]: if there exists a road, there exists a sidewalk.

Dilation is a mathematical morphology operation that expands or dilates the shapes present in a binary input image. It operates on an image f using a structuring element g. The structuring element is translated across the image pixel by pixel. When the intersection between g and f is non-empty, the corresponding pixel is considered valid. As a result, dilation produces the set of pixels where the structuring element overlaps with at least one pixel of the image.

Formally, morphological dilation can be expressed as follows. Let $p = (p_1, p_2)$ denote the corresponding pixel, and let g_p denote the translation of the structuring element g with its origin placed at p. The dilation of f by g is then given by:

$$\delta_g(f) = \{\, p \mid g_p \cap f \neq \emptyset \,\}. \quad (1)$$

Equivalently, dilation produces the collection of pixels p where the shifted structuring element g_p has a non-empty intersection with f [22].

This technique is particularly effective for connecting or approximating nearby features, such as road surfaces with adjacent sidewalks. In this work, dilation was implemented using the OpenCV Python library. A square structuring element of size 15×15 pixels was employed, created with `np.ones((15,15), np.uint8)`. The dilation was applied with a single iteration using `cv2.dilate(f, g, iterations=1)`. Preliminary tests with smaller kernels (e.g., 10×10) did not sufficiently capture sidewalk adjacency, while larger kernels (e.g., 20×20) produced excessively wide sidewalks. The 15×15 kernel provided the best balance between sidewalk expansion and spatial consistency with the ground truth.

D. Evaluation Metric

To evaluate segmentation performance, we employed the Jaccard Index, also known as Intersection over Union (IoU) [23]. IoU is a widely used metric in semantic segmentation tasks due to its scale-invariant properties. It is calculated as the ratio between the overlapping area of the predicted segmentation and the ground truth, divided by the total combined area of both. The score ranges from 0 to 1, where 1 indicates a perfect overlap. A visual explanation of this metric is provided in Figure 2.

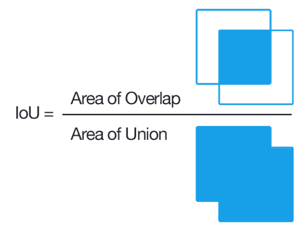

Fig. 2. Illustration of Intersection over Union (IoU). The numerator represents the intersection of pixels between the predicted and ground truth objects, while the denominator represents their union.

E. Digital Tools

The following tools were used throughout the study:

- *Python:* the primary programming language used for data processing, model execution, and automation. Running Tile2Net requires a machine equipped with a GPU (Graphics Processing Unit) to efficiently process large volumes of image tiles.
- *R:* the tool used for raster and vector data processing, NDWI calculation, and geospatial visualization tasks. All operations were performed with version 4.5.0 of the R language [24].
- *Geosampa:* an open urban data platform used to obtain high-resolution aerial images and associated tabular datasets for the city of São Paulo [10].
- *QGIS (Quantum Geographic Information System):* a free and open-source GIS software used to preprocess and divide GeoSampa aerial images into the tile format required for model input [25].

III. Experiments, Results and Discussion

The borough of São Miguel covers a total area of 24.3 km² and has a population of approximately 369,496 residents [10], [26]. Data for its three constituent districts are summarized in Table II.

As a pilot test, aerial images from a selected region of interest (ROI) within São Miguel were used. A zoomed-in view of this region is presented in Figure 3, revealing buildings, roads, and sidewalks, some of which are partially obscured by trees and shadows. Subsequently, the high-resolution aerial imagery of the ROI is shown in the upper-left panel of Figure 4. Tiles

TABLE II
SOCIO-SPATIAL AND DEMOGRAPHIC CHARACTERISTICS OF THE BOROUGH OF SÃO MIGUEL (SÃO PAULO, BRAZIL). FONTS: [10], [26]

Administrative division	
Borough	São Miguel
Districts	São Miguel, Jardim Helena, Vila Jacuí
Geographic location	
Zone	East Zone (Zona Leste)
Coordinates (approx. centroid)	23°30'10"S 46°26'21"W
Area and density	
Total area	24.3km²
Population density	15,202.3 inhabitants/km²
Social vulnerability	
São Paulo State Social Vulnerability Index (IPVS)	25.6%
Municipal Human Development Index (MHDI)	0.736

Fig. 3. Zoom in of the aerial image of the borough of São Miguel used as a pilot test region.

were rendered at zoom level 19, with the ROI bounded from SW (-46.44607, -23.50607) to NE (-46.43920, -23.49977).

A. Results for Sidewalk Semantic Segmentation

Figure 4 upper right presents the semantic segmentation results overlaid on the aerial image of the ROI, as produced by the Tile2Net tool. In this visualization, sidewalks are represented in blue, roads in green, and crosswalks in red. The primary focus of this study is the accurate segmentation of sidewalks.

Although road segmentation is not the central objective, the green road mask offers a valuable input for post-processing operations, such as morphological dilation, illustrated later in this section.

A comparison between the sidewalk segmentation output from Tile2Net (Figure 4 bottom left) and the manually annotated ground truth resulted in an Intersection over Union (IoU) score of 9.5%.

It is important to note that the current version of Tile2Net was not fine-tuned for the urban fabric of the city of São Paulo. The original model was trained on aerial images from cities in the United States, where it achieved a sidewalk IoU of 82.67% on its test set [8].

To improve segmentation performance in our case study, a simple dilation operation was applied to the road mask (Figure 4 bottom right), which increased the IoU to 39%. While this result represents more than a threefold improvement, it still falls short of the performance achieved in the U.S. context, highlighting the limitations of direct model transferability across geographic regions, and the importance of fine-tuning the model with the aerial images of the city of São Paulo.

B. Seasonal Variation Analysis of Flood-Prone Areas

The comparative analysis of the Normalized Difference Water Index (NDWI) revealed a significant seasonal variation in surface water coverage (Figure 5).

Results showed a 60.54% reduction in water-covered areas between the rainy season (February) and dry season (July), with absolute values decreasing from 1.47% (0.39 km^2) to 0.58% (0.15 km^2) of the total investigated area. This 0.89 percentage point variation highlights the notable local hydrological susceptibility, with seasonal flooding patterns predominantly concentrated along the Tietê River and the Itaim and Lajeado streams.

The spatial interface between flood-prone areas and the urban grid therefore constitutes an essential factor for developing resilient pedestrian infrastructure and identifying safe evacuation routes during extreme events. The episodically recurrent nature of these flooding phenomena demands adapted urban planning solutions, particularly in contexts marked by socio-spatial vulnerabilities, as observed in São Miguel.

Sentinel-2 satellite imagery, with 10-meter spatial resolution, proved suitable for detecting seasonal variability in surface waters. Its high temporal resolution enables continuous monitoring across different hydrological seasons, supporting both short-term assessments and long-term resilience strategies.

IV. CONCLUSION AND FUTURE WORKS

This ongoing study explored the semantic segmentation of sidewalks using the pre-trained Tile2Net model, applied to aerial imagery of the São Miguel district, São Paulo. Although the model was originally trained on imagery from U.S. cities, its application in the local urban context demonstrated potential, especially when combined with simple post-processing techniques, such as morphological dilation, which improved

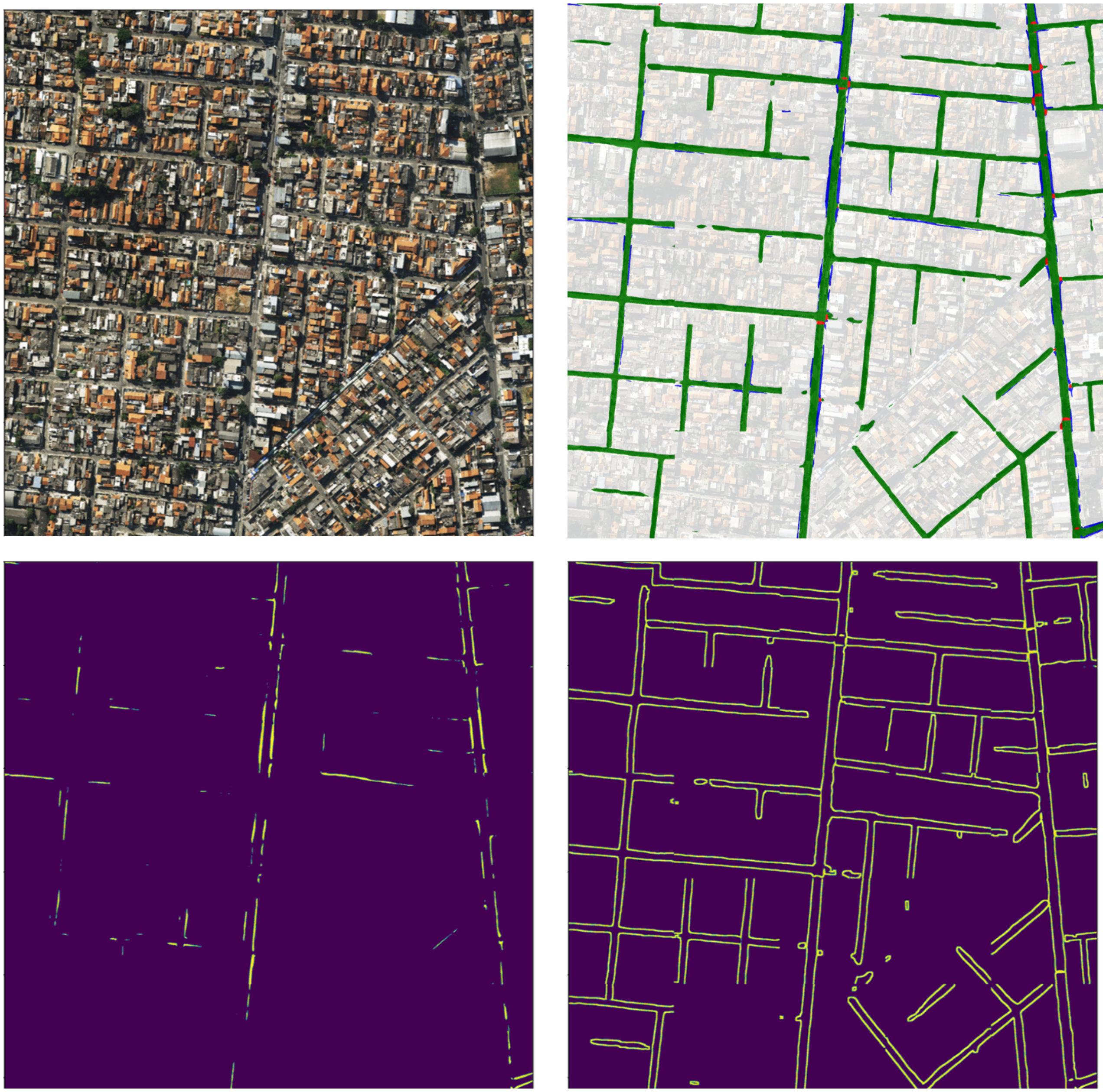

Fig. 4. Upper left: aerial image of the borough of São Miguel, used as the pilot test area. Upper right: semantic segmentation mask overlaid on the aerial image, with sidewalks in blue, roads in green, and crosswalks in red. Bottom left: sidewalk segmentation mask generated by the Tile2Net model. Bottom right: sidewalk segmentation mask after applying dilation to the road mask obtained from the Tile2Net output (roads shown in green in the upper right image).

the consistency of the extracted masks. In parallel, hydrological analysis using the NDWI index enabled the identification of seasonal water masks, highlighting the vulnerability of the territory to flooding and inundation events, particularly along the Tietê River and nearby streams.

To date, the consolidation of the multimodal dataset has included: (i) acquisition of Sentinel-1, Sentinel-2, and aerial imagery, along with the collection of socioeconomic information; (ii) pre-processing of Sentinel-2 images and orthophotos; (iii) preliminary application and evaluation of the Tile2Net model for sidewalk segmentation in São Miguel; (iv) implementation of morphological post-processing techniques to refine the generated masks; and (v) seasonal analysis of flooded areas using the NDWI index.

The next steps involve fine-tuning the Tile2Net model with annotated aerial imagery from São Paulo in order to enhance its performance across different urban scenarios. Additionally, the generated masks will be further refined using advanced post-processing techniques such as inpainting and geometric shape completion, followed by the conversion of segmented sidewalks into graphs for pedestrian connectivity analysis and their integration with flood masks. Complementarily, the use of Sentinel-1 SAR data will be explored with the aim of capturing flood dynamics across different temporal and climatic scenar-

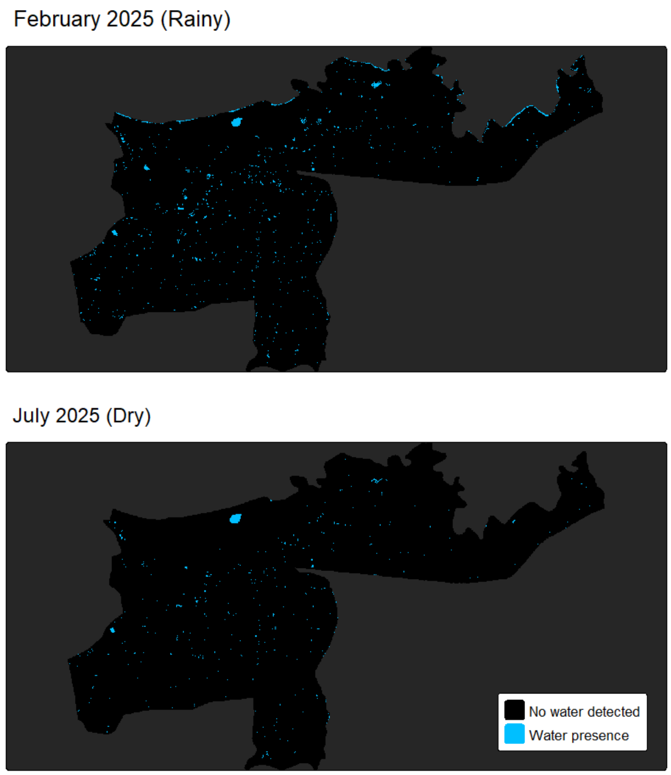

Fig. 5. NDWI derived from Sentinel-2 imagery (10 m), comparing a rainy day (February) and a dry day (July) in São Miguel, SP.

ios. In subsequent stages, additional areas beyond São Miguel will be tested, enabling the evaluation of the methodology in distinct socio-spatial and environmental contexts within the city.

Thus, the multimodal dataset currently under development represents a promising foundation for advancing computational methodologies, fostering climate-risk-sensitive urban planning, and supporting applications aimed at enhancing the resilience, livability, and sustainable planning of cities.

V. Acknowledgment

The authors are grateful to FAPESP grant #2022/15304-4, #2024/17415-3, #2025/02274-8, CNPq, CAPES, FINEP and MCTI PPI-SOFTEX (TIC 13 DOU 01245.010222/2022-44).

References

[1] M. Rajaee, B. Echeverri, Z. Zuchowicz, K. Wiltfang, and J. Lucarelli, "Socioeconomic and racial disparities of sidewalk quality in a traditional rust belt city," *SSM-Population Health*, 2021.

[2] K. Lowe, "Environmental justice and pedestrianism: Sidewalk continuity, race, and poverty in new orleans, louisiana," *Transportation Research Record*, vol. 2598, no. 2598, pp. 119–123, 2016.

[3] R. J. Lores, "O vexame das calçadas," [Online]. Available: https://www.youtube.com/watch?v=81rZwFtq9Qk&list=FLbxTiNw3s9f2Uj7ZOfax8gQ, 2024, accessed: 2025-07-27.

[4] P. J. Gibson and W. E. Marshall, "Disparate approaches to maintaining roads and sidewalks: An interview study of 16 u.s. cities," *Transportation Research Record*, vol. 2676, no. 9, pp. 553–567, 2022.

[5] R. Stefanidis and A. Bartzokas-Tsiompras, "Pedestrian accessibility analysis of sidewalk-specific networks: Insights from three latin american central squares," *Sustainability*, 2024.

[6] K. A. Shastry, B. Jagadish, V. Bhat, and A. Chowdary, "Automated urban infrastructure classification using aerial imagery," in *2024 International Conference on Advances in Computing and Information Sciences (IACIS)*, 2024.

[7] S. Liang, A. Guo, B. Liang, Z. Li, Y. Ding, Y. Wang, and F. Chen, *Machine Learning and Computer Vision Applications in Civil Infrastructure Inspection and Monitoring*. John Wiley Sons, Ltd, 2024, ch. 4, pp. 59–80.

[8] M. Hosseini, A. Sevtsuk, F. Miranda, R. M. C. Junior, and C. T. Silva, "Mapping the walk: A scalable computer vision approach for generating sidewalk network datasets from aerial imagery," *Computers, Environment and Urban Systems*, vol. 101, p. 101950, 2023.

[9] V. Bonets and C. Figueiredo, "Vídeos: Após chuvas, ruas ficam debaixo d'água na zona leste de são paulo," [Online]. Available: https://www.cnnbrasil.com.br/nacional/sudeste/sp/videos-apos-chuvas-ruas-ficam-debaixo-dagua-na-zona-leste-de-sao-paulo, 2025, accessed: 2025-07-20.

[10] Prefeitura de São Paulo, *Mapa Digital da Cidade de São Paulo*, https://geosampa.prefeitura.sp.gov.br/PaginasPublicas/_SBC.aspx, 2025, accessed: 2025-07-20.

[11] ——, *São Miguel Paulista*, https://prefeitura.sp.gov.br/web/sao_miguel_paulista/w/historico/432, 2025, accessed: 2025-07-19.

[12] MapTiler Team, *Coordinate Systems Worldwide*, https://epsg.io/, 2025, accessed: 2025-07-20.

[13] European Space Agency (ESA), "Introducing sentinel-2," https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Introducing_Sentinel-2, 2025, accessed: Jul. 20, 2025.

[14] B.-C. Gao, "NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sensing of Environment*, vol. 58, no. 3, pp. 257–266, 1996.

[15] A. Tao, K. Sapra, and B. Catanzaro, "Hierarchical multi-scale attention for semantic segmentation," 2020. [Online]. Available: https://arxiv.org/abs/2005.10821

[16] L. van Oorschot, D. Michailidis, N. IJzerman, S. Jansen, and D. Roijers, "Where (not) to cross the street," https://bnaic2024.sites.uu.nl/wp-content/uploads/sites/986/2024/11/Where-Not-to-Cross-the-Street.pdf, 2024.

[17] M. Saberi, T. Lilasathapornkit, F. Nourmohammadi, A. Emami, and M. Damandeh, "Mapping, modeling, and measuring walking and walkability: The role of ai in enhancing urban pedestrian networks," https://accessiblecities.github.io/UrbanAccess2024/assets/papers/Saberi_MappingModelingAndMeasuringWalkingAndWalkability-TheRoleOfAIInEnhancingUrbanPedestrianNetworks_UrbanAccess2024.pdf, 2024.

[18] Y. Zhang, B. Howe, and A. Caspi, "Reliable, routable, and reproducible: Collection of pedestrian pathways at statewide scale," 2024. [Online]. Available: https://arxiv.org/abs/2410.19762

[19] Y. Zhang, B. Howe, S. Mehta, N.-J. Bolten, and A. Caspi, "Pathwaybench: Assessing routability of pedestrian pathway networks inferred from multi-city imagery," 2024. [Online]. Available: https://arxiv.org/abs/2407.16875

[20] R. I. Sultan, C. Li, H. Zhu, P. Khanduri, M. Brocanelli, and D. Zhu, "GeoSAM: Fine-tuning sam with multi-modal prompts for mobility infrastructure segmentation," 2024. [Online]. Available: https://arxiv.org/abs/2311.11319

[21] T. Senlet and A. Elgammal, "Satellite image based precise robot localization on sidewalks," in *2012 IEEE International Conference on Robotics and Automation*. IEEE, 2012, pp. 2647—-2653.

[22] L. da Fontoura Costa and R. M. C. Junior, *Shape Analysis and Classification: Theory and Practice*. CRC Press, Dec. 2010. [Online]. Available: https://doi.org/10.1201/9781420037555

[23] D. W. Goodall, "A new similarity index based on probability," *Biometrics*, vol. 22, pp. 882—-907, 1966. [Online]. Available: https://doi.org/10.2307/2528080

[24] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, 2025, (Version 4.5.0) [Computer software]. [Online]. Available: https://www.r-project.org/

[25] QGIS Project, *QGIS Official Site*, 2025, accessed on 2025-07-27. [Online]. Available: https://www.qgis.org/en/site/

[26] Governo Brasileiro, *Instituto Brasileiro de Geografia e Estatística (IBGE)*, https://www.ibge.gov.br/, 2025, accessed: 2025-07-20.