



UNIVERSIDADE  
ESTADUAL DE LONDRINA

---

ANA LUIZA FAVARÃO LEÃO

**MULTISCALED WALKABILITY:**

Exploring Built Environment using artificial intelligence.

---

Londrina  
2023

ANA LUIZA FAVARÃO LEÃO

**MULTISCALED WALKABILITY:**

Exploring Built Environment using artificial intelligence.

Thesis dissertation submitted for the partial fulfillment of the requirements for the Academic Degree of Doctor of Architecture and Urbanism from the Associated Program in Architecture and Urbanism from the State University of Londrina and State University of Maringá.

Londrina  
2023

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

Favarão Leão, Ana Luiza .

Multiscaled walkability: : Exploring Built Environment using artificial intelligence / Ana Luiza Favarão Leão. - Londrina, 2023.  
138 f. : il.

Orientador: Milena Kanashiro.

Tese (Doutorado em Arquitetura e Urbanismo) - Universidade Estadual de Londrina, Centro de Tecnologia e Urbanismo, Programa de Pós-Graduação em Arquitetura e Urbanismo, 2023.

Inclui bibliografia.

1. Caminhabilidade - Tese. 2. Ambiente Construído - Tese. 3. Aprendizado Profundo - Tese. 4. Google Street View - Tese. I. Kanashiro, Milena. II. Universidade Estadual de Londrina. Centro de Tecnologia e Urbanismo. Programa de Pós-Graduação em Arquitetura e Urbanismo. III. Título.

CDU 711/72

ANA LUIZA FAVARÃO LEÃO

**MULTISCALED WALKABILITY:**

**Exploring Built Environment using artificial intelligence.**

Thesis dissertation presented to the State University of Londrina – UEL, as a partial requirement for obtaining the title of Doctor in Architecture and Urbanism

Advisor: Profa. Dra. Milena Kanashiro

---

Profa. Dra. Milena Kanashiro  
Universidade Estadual de Londrina - UEL

---

Prof. Dr. Rodrigo Siqueira Reis  
Washington University in St. Louis - WashU

---

Prof. Dr. Júlio Celso Borello Vargas  
Universidade Federal do Rio Grande do Sul – UFRGS

---

Prof. Dr. Renato Leão Rego  
Universidade Estadual de Maringá – UEM

---

Profa. Dra. Ercilia Hirota  
Universidade Estadual de Londrina - UEL

Londrina, September 19<sup>th</sup>, 2023.

## AGRADECIMENTOS

Agradeço ao povo brasileiro por financiar minha formação de maneira pública e gratuita. À Universidade Estadual de Londrina, sou grata por proporcionar uma formação de excelência desde a graduação até o doutorado. Agradeço também à Comissão Fulbright no Brasil e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

A todos os integrantes dos grupos de pesquisa Design Ambiental Urbano e Cidade: Movimento e Atividades, minha sincera gratidão. Vocês foram parceiros fundamentais ao longo de todos estes anos, nas pesquisas, reuniões e trabalhos realizados. Acredito que nossa união e os resultados alcançados como grupo só foram possíveis graças à verdade, empatia e companheirismo que compartilhamos. Agradeço aos alunos e orientados que me permitiram aprender e ensinar durante estes anos de doutoramento.

Agradeço imensamente a todos os mestres e orientadores que estiveram comigo. À Prof<sup>ª</sup>. Dra. Milena Kanashiro expresso minha mais profunda gratidão por ter me acolhido como sua aluna e por ter me proporcionado, ao longo destes anos, oportunidades para aprender com você e crescer como pesquisadora. Agradeço por sempre ver em mim potenciais além dos que eu mesma percebia, e por toda a paciência e confiança que sempre depositou em mim. Sigo com a certeza de que a profissional que me tornei deve-se à sua orientação e às oportunidades que me deu. Agradeço, sobretudo, pela amizade, carinho e parceria.

Ao Prof. Dr. Rodrigo Siqueira Reis, agradeço por me acolher em St. Louis e fazer-me sentir sempre bem-vinda. Em um breve período, compartilhou comigo ensinamentos que guardarei para o resto da minha vida. Sua sabedoria, generosidade e sensibilidade me edificaram e mudaram como pesquisadora e pessoa.

Aos membros da banca avaliadora desta tese de doutorado, Prof. Dr. Renato Leão, Prof. Dr. Júlio Vargas e Prof<sup>ª</sup>. Ercília Hirota, expresso minha sincera gratidão. Tenho profunda estima e admiração por cada um de vocês e agradeço o tempo dedicado a contribuir com a minha formação.

Agradeço minha família, em especial as mulheres que me guiam – minha mãe e avó, Alda e Aparecida. Vocês são minhas raízes, e com essa convicção, sinto-me fortalecida para enfrentar cada nova fase sem me perder de mim mesma. Com vocês, aprendi a valorizar o presente sem medo dos desafios, porque somos feitas para nos adaptar e conquistar nossos sonhos. Amo-as de todo o meu coração.

Expresso minha gratidão ao meu companheiro de vida, Hugo, que na sua serenidade e mansidão é também imensidão. Amar você e compartilhar desta trajetória são um imenso privilégio. Obrigada por todo o apoio que sempre me proporcionou e por acreditar em mim, mesmo quando eu duvidava de minha própria capacidade.

Por fim, expresso minha gratidão a todos os amigos que, de alguma maneira, fizeram parte e contribuíram para esta trajetória de doutoramento. Muito obrigada a todos!

## RESUMO

*Caminhar é reconhecido há muito tempo na literatura científica como o modo de transporte mais equitativo e universalmente acessível nas cidades. É uma forma predominante de atividade física para adultos. A caminhabilidade descreve o grau em que um ambiente apoia e incentiva a caminhada. Tradicionalmente, esse atributo urbano tem sido quantificado usando índices, que são essencialmente equações ponderadas que encapsulam vários fatores de bairro em macroescala. No entanto, essas métricas muitas vezes negligenciam características de rua em microescala que afetam significativamente a caminhada, principalmente devido aos desafios da coleta de dados. Existe uma crescente necessidade de pesquisa sobre caminhabilidade em cidades brasileiras de médio a pequeno porte, dada a rápida crescimento populacional, disparidades sociais e a importância da caminhada como modo de transporte. Este estudo visa elucidar sobre a caminhabilidade em múltiplas escalas, através do desenvolvimento de modelos de caminhabilidade para cidades medias e pequenas do Brasil usando técnicas de aprendizado profundo. A tese central postula que uma análise robusta de caminhabilidade em nível de bairro, para cidades brasileiras de médio a pequeno porte, pode ser desenvolvida usando técnicas de aprendizado profundo. Utilizando a metodologia de Pesquisa em Ciência do Design, três cidades de tamanhos variados no Paraná, Brasil - Rolândia, Cambé e Londrina - foram selecionadas como estudos de caso. Fatores de caminhabilidade em macroescala foram organizados em conjunto com características em microescala, extraídas do Google Street View usando um modelo de rede neural profunda para segmentação semântica. Os dados dessas cidades foram então avaliados usando três modelos de regressão multinível distintos: um focando apenas em fatores em microescala, outro em fatores em macroescala, e um modelo compreensivo em múltiplas escalas abrangendo variáveis de ambas as escalas. A variável de resposta primária foi o modo de transporte - caminhar versus outros - com variáveis sociodemográficas consideradas como covariáveis. Esses dados foram obtidos do Plano de Mobilidade das respectivas cidades. As descobertas revelaram padrões variados nas três cidades. Em Rolândia, a menor cidade analisada, fatores como a presença de muros e indivíduos em bicicletas ou motocicletas foram vistos como promotores da caminhada. Por outro lado, sinalizações urbanas, como postes e semáforos, desencorajavam a caminhada. Em Cambé, de tamanho intermediário, a presença de terrenos vagos e espaços verdes estava associada a uma redução na caminhada. Em Londrina, a maior cidade considerada neste estudo, uma combinação de fatores micro e macro influenciou o comportamento da caminhada. A disponibilidade de calçadas e o uso diversificado do uso do solo incentivaram a caminhada, enquanto uma maior presença de estradas e áreas verdes estava negativamente associada à caminhada. Esses resultados ampliam nosso entendimento sobre caminhabilidade e podem informar futuros esforços de planejamento urbano. Ao focar nessas abordagens, os formuladores de políticas podem elaborar políticas urbanas específicas ao contexto para promover a caminhada, combinando estratégias baseadas em evidências.*

**Palavras-chave:** Caminhabilidade. Ambiente Construído. Aprendizado Profundo. Google Street View.

## ABSTRACT

*Walking has long been recognized in scientific literature as the most equitable and universally accessible mode of transportation within cities. It is a predominant form of physical activity for adults. Walkability describes the degree to which an environment supports and encourages walking. Traditionally, this urban attribute has been quantified using indices, which are essentially weighted equations that encapsulate various macro-level neighborhood factors. However, these metrics often overlook micro-level street characteristics that significantly affect walking, primarily due to the challenges of data collection. There is a growing need for walkability research in medium to small Brazilian towns, given the rapid population growth, social disparities, and the importance of walking as a mode of transportation. This study aims to shed light on multi-scale walkability by devising walkability models for such towns in Brazil. The central thesis posits that a robust multi-scale, neighborhood-level walkability analysis for medium-small Brazilian towns can be developed using deep learning techniques. Employing the Design Science Research methodology, three towns of varying sizes in Paraná, Brazil—Rolândia, Cambé, and Londrina—were selected as case studies. Macro-scale walkability factors were organized in conjunction with microscale features, extracted from Google Street View using a semantic segmentation deep neural network model. The data from these cities was then evaluated using three distinct multilevel regression models: one focusing solely on micro-scale factors, another on macro-scale factors, and a comprehensive multi-scale model encompassing variables from both scales. The primary response variable was the mode of transport—walking versus others—with sociodemographic variables factored in as covariates. This data was sourced from the Mobility Plan of the respective towns. The findings revealed varied patterns across the three cities. In Rolândia, factors like the presence of walls and individuals on bikes or motorcycles were seen to promote walking. Conversely, urban signage, such as posts and traffic lights, deterred walking. In Cambé, which has approximately 100,000 residents, the presence of vacant lots and green spaces corresponded to reduced walking. In Londrina, the largest town in this study, a mix of micro and macro factors influenced walking behavior. The availability of sidewalks and diverse land use encouraged walking, while an increased prevalence of roads and green spaces discouraged it. These results enhance our understanding of walkability and can inform future urban planning efforts. By focusing on these insights, policymakers can craft context-specific urban policies to promote walking, combining evidence-based strategies with local nuances.*

**Keywords:** Walkability. Built Environment. Deep Learning. Google Street View.

## FIGURE INDEX

|   |    |
|---|----|
| Figure 1 – Research Outline. ....   | 8  |
| Figure 2 - Image Processing Example Based on Semantic Segmentation. ....  | 27 |
| Figure 3– Country distribution of research records obtained from literature search (n=45)....   | 30 |
| Figure 4– Yearly distribution of research records obtained from literature search (n=46).....   | 31 |
| Figure 5 - Development of the research process in three phases.....   | 34 |
| Figure 6 – Case study location, Rolândia, Cambé and Londrina -Brazil. ....  | 35 |
| Figure 7 - Aggregation units for the three case studies – census tracts and zones.....  | 37 |
| Figure 8 - Methodological workflow for macro-scale Walkability.....   | 39 |
| Figure 9 - Methodological workflow for Micro-scale Walkability. ....  | 45 |
| Figure 10 - Example of the panoramic image montage from each geo-tagged point. ....   | 46 |
| Figure 11 - Representation of the aggregate absolute number of walking .....  | 55 |
| Figure 12 – Histogram of walking distances of the n=394 walking trips .....   | 56 |
| Figure 13 - Macro variable for Rolândia. Respectively: residential parcels, built area (building footprint) for commercial purposes, intersections, land use data, street integration calculations from space syntax. ....                                    | 57 |
| Figure 14- Sampling of geo-tagged points for image extraction in Rolândia. ....   | 60 |
| Figure 15 - Spearman’s rho correlation between absolute walking trips per census tract and identified category values per zone – Rolândia.....  | 61 |
| Figure 16 - Boxplot representing the importance of micro-scaled categories.....   | 62 |
| Figure 17 – Hotspot analysis for the distribution of the Bike, Pole, Rider and Wall categories for Rolândia. ....   | 65 |
| Figure 18 - Representation of the aggregate absolute number of walking trips conducted by census tracts in Cambé. ....  | 67 |
| Figure 19 – Histogram of walking distances of the n=1080 walking trips conducted in Cambé. ....   | 68 |
| Figure 20 - Data used for measuring variables at the macro scale for Cambé. Respectively: residential parcels, built area (building footprint) for commercial purposes, intersections, land use data, street integration calculations from space syntax. .... | 69 |
| Figure 21 - Sampling of geo-tagged points for image extraction in Cambé. ....   | 72 |
| Figure 22 - Spearman’s rho correlation between absolute walking trips per zone and identified category values per zone – Cambé .....  | 73 |

|   |    |
|---|----|
| Figure 23 - Boxplot representing the importance of micro-scaled categories obtained from semantic segmentation to walking (or not) in Cambé.....  | 74 |
| Figure 24 - Hot Spot for the distribution of the Vegetation in Cambé.....   | 77 |
| Figure 25 - Hot Spot for the distribution of the Terrain category in Cambé.....   | 77 |
| Figure 26 - Representation of the aggregate absolute number of.....   | 80 |
| Figure 27 – Histogram of walking distances of the n=4162 walking trips conducted in Londrina.<br>.....  | 81 |
| Figure 28 - Residential and commercial lots and their building projections from Londrina. ..  | 82 |
| Figure 29 -True Intersections and land use data from Londrina. ....   | 82 |
| Figure 30 - Street Integration at the 1200-meter radius.....  | 83 |
| Figure 31 - Sampling of geo-tagged points for image extraction.....   | 87 |
| Figure 32 - Spearman’s rho correlation between absolute walking trips per zone and identified category values per zone in Londrina.....           | 89 |
| Figure 33 - Boxplot representing the importance of micro-scaled categories obtained from semantic segmentation to walking (or not). Londrina..... | 90 |
| Figure 34 - Hot Spot for the distribution of the Person, Road, Sidewalk and Vegetation categories in Londrina. ....                               | 94 |

## TABLE INDEX

|  |     |
|--|-----|
| Table 1 - Environmental characteristics and relationships to walking behavior.....   | 25  |
| Table 2 - Summary of most common environmental features in micro-scale audits.....   | 23  |
| Table 3 – Sample sizes for OD survey data, adopted sociodemographic variables, their classes, and data sources.....  | 38  |
| Table 4 – Macroscale walkability variables considered and data sources. ....   | 40  |
| Table 5 – Land use category modifications for Londrina. ....   | 41  |
| Table 6 - List of annotated classes in the Cityscapes dataset. ....  | 48  |
| Table 7 - Sociodemographic characteristics of the sample for Rolândia (n = 2152). ....   | 53  |
| Table 8- Sociodemographic characteristics of the sample versus walking or not in Rolândia (n = 2152).....  | 54  |
| Table 9 - Multilevel logistic regression analysis for the relationship between the choice of walking, macro/macro-built environment, and individual factors (n = 2152).....                  | 59  |
| Table 10 - Multilevel logistic regression analysis on walking, individual factors and micro-level-built environment characteristics for Rolândia (n=2152).....                               | 64  |
| Table 11 - Sociodemographic characteristics of the sample .....  | 66  |
| Table 12 - Sociodemographic characteristics of the sample versus .....   | 67  |
| Table 13 - Multilevel logistic regression analysis for the relationship between the choice of walking, macro/macro-built environment, and individual factors for Cambé (n = 3348).....       | 71  |
| Table 14 - Multilevel logistic regression analysis for the relationship between walking or not, individual factors and micro-level-built environment characteristics for Cambé (n=3348). ... | 76  |
| Table 15 - Sociodemographic characteristics of the sample. (n = 18635). ....   | 78  |
| Table 16- Sociodemographic characteristics of the sample versus walking or not in Londrina (n = 18635). ....   | 79  |
| Table 17 - Multilevel logistic regression analysis for the relationship between the choice of walking, built environment, and individual factors (n = 18635). ....                           | 84  |
| Table 18 - Multilevel logistic regression analysis for the relationship between the choice of walking, micro scale-built environment, and individual factors (n = 18635). ....               | 92  |
| Table 19 - Multilevel logistic regression analysis for the relationship between the choice of walking, .....   | 99  |
| Table 20 – Percentage classes from all obtained GSV images for all cities. ....  | 107 |

## LIST OF ABBREVIATIONS

|       |   |
|-------|---|
| AIC   | Akaike Information Criteria                     |
| BE    | Built Environment                               |
| CI    | 95% confidence intervals                        |
| CNN   | Convolutional Neural network                    |
| DSR   | Design Science Research                         |
| FAR   | Floor Area Ratio                                |
| GSV   | Google street view                              |
| IBGE  | Instituto Brasileiro de Geografia e Estatística |
| ICC   | The Interclass Correlation Coefficient          |
| IPPUL | Institute for Research and Urban Planning       |
| mIOU  | Mean pixel intersection-over-union              |
| MW    | Minimum wage                                    |
| OD    | Origin-Destination                              |
| OR    | Odds Ratio                                      |
| OSM   | OpenStreetMap                                   |
| SVI   | Street View Imagery                             |
| UEL   | Universidade Estadual de Londrina               |
| UEM   | Universidade Estadual de Maringá                |
| VIF   | Variance Inflation Factor                       |
| WI    | Walkability indices                             |
| PNMU  | National Urban Mobility Policy                  |

## TABLE OF CONTENTS

|  |            |
|--|------------|
| <b>TABLE OF CONTENTS .....</b>   | <b>1</b>   |
| <b>1 INTRODUCTION .....</b>  | <b>1</b>   |
| 1.1 Thesis and research question   | 4          |
| 1.2 Research Objectives  | 5          |
| 1.3 Research Outline and Thesis Structure  | 5          |
| <b>2 BIBLIOGRAPHIC REVIEW.....</b>   | <b>21</b>  |
| 2.1 The macro scale approach to Walkability  | 23         |
| 2.2 The micro scale approach to Walkability  | 21         |
| 2.3 Machine learning applied to urban analyses   | 24         |
| 2.3.1 Deep learning using Street View Images for Walkability: state of art                       | 27         |
| <b>3 METHODOLOGY AND DATA .....</b>  | <b>33</b>  |
| 3.1 Case studies: Walking data and aggregation.  | 34         |
| 3.2 The first methodological phase – Macro scale Walkability                                     | 38         |
| 3.3 Second Phase Microscale Walkability  | 45         |
| 3.4 Third Phase Comprehensive Walkability modeling   | 50         |
| <b>4 RESULTS AND DISCUSSIONS .....</b>   | <b>53</b>  |
| <b>4.1 RESULTS FOR ROLÂNDIA .....</b>  | <b>53</b>  |
| 4.1.1 Macroscale Walkability for Rolândia  | 56         |
| 4.1.2 Microscale Walkability for Rolândia  | 59         |
| <b>4.2 RESULTS FOR CAMBÉ .....</b>   | <b>66</b>  |
| 4.2.1 Macroscale Walkability for Cambé   | 68         |
| 4.2.2 Microscale Walkability for Cambé   | 71         |
| <b>4.3 RESULTS FOR LONDRINA .....</b>  | <b>78</b>  |
| 4.3.1 Macroscale Walkability for Londrina  | 81         |
| 4.3.2 Microscale Walkability for Londrina  | 86         |
| <b>4.4 RESULTS AND DISCUSSIONS OF MULTISCALE MODELLING IN ROLÂNDIA, CAMBÉ AND LONDRINA .....</b> | <b>95</b>  |
| <b>4.5 EVALUATION OF THE DESIGN PROCESS AND DISCUSSIONS OF MULTISCALE WALKABILITY.....</b>       | <b>104</b> |
| <b>5 CONCLUSIONS.....</b>  | <b>112</b> |

## 1 INTRODUCTION

Walking has been historically understood in the scientific literature as the most equitable and accessible mean of moving within the city. It represents the most common form of adults' physical activity (Orellana et al., 2016). Further, it contributes to mental wellbeing and can enrich communities through social capital (Leyden, 2003). Evidence indicates that individuals who walk more not only tend to present an average body mass index (Booth et al., 2005) but are also reported to have generally lower stress levels (Barton et al., 2009; Sattler et al., 2018). Moreover, in a more ample perspective, the shift from car driving to walking reduces air pollutant emissions (Watts et al., 2015).

In this context, the built environment can either facilitate or hinder walking behaviors (B. Saelens & Handy, 2008), and it's considered a significant determinant of walking for enhancing both population and planetary health (Dean et al., 2020a). Areas that promote walking typically possess distinct features that render them walkable (Saelens; Handy, 2008). Walkability, as primarily defined by Leslie and Cerin (2008), is "the extent to which characteristics of the built environment and land use may encourage or discourage walking for leisure, exercise, recreation, service access, or commuting to work." Recent discussions highlight the complexity of walkability, suggesting that it is a multifaceted concept that could benefit from a broader definition to encapsulate the environmental aspects linked to promoting physical activity (Tobin et al., 2022). Therefore, in simpler terms, walkability represents how conducive the built environment is for walking (Wesemael, 2020).

Cervero and colleagues (Cervero & Kockelman, 1997; R. Ewing & Cervero, 2010) have provided the theoretical foundation for measuring walkability, dividing the built environment into different dimensions, i.e., density, diversity, design, distance, destination, and demand management. Through these ideas, walkability indices have been developed to predict inhabitants' walking activity and active travel at the neighborhood level (H. Wang & Yang, 2019a). Indices sum variables in weighted equations, representing many factors that compose walkability. This proposition's idea is that singular variables are not as robust as indices to represent such a complex urban quality.

Since the seminal work by Frank et al. (2009), which forwarded one of the most largely accepted walkability indices (WI) in the literature, many others have been proposed to objectively quantify built environment's attributes that may impact walking (Cervero & Kockelman, 1997; L. D. Frank et al., 2009; Lotfi & Koohsari, 2011; Lovasi et al., 2008; Owen et al., 2007; Reis et al., 2013; Van Cauwenberg et al., 2016). These metrics consider 'macro'

neighborhood features (C. Harvey & Aultman-Hall, 2016), such as residential density, land use mixture (schools, shops, etc.), and street connectivity, in collective assessment (Maghelal & Capp, 2017). Such indices represent a clear literature tendency in evaluating walkability (Arellana et al., 2019), understanding it as a product of all built environment attributes that act as predictors of health (Shashank & Schuurman, 2019).

On the other hand, ‘macro’ focused walkability indices have been criticized over recent years, for (a) might not being suited for different social, cultural, and urban realities (Leão, 2019; H. Wang & Yang, 2019a) (b) for their inherent assumptions in the weighting of variables in the equations, that seem to be theoretically cloudy (Shashank & Schuurman, 2019) and mainly (c) not capturing street-level (micro-scale) built environment characteristics, such as the presence of trees, the quality and width of sidewalks, and the quality of streets (Kim et al., 2014), variables that play an essential role in walkability (Larranaga et al., 2018). A macro-scale perspective emphasizes the overarching dimensions and patterns, illustrating how various streetscapes interrelate when observed from an aerial view. However, it doesn't delve into the specific compositions of streetscapes, be it in terms of larger elements like buildings and trees or finer details such as architectural designs and materials (C. Harvey & Aultman-Hall, 2016).

Street-level characteristics, microwalkability (Park et al., 2014) or micro-scale, are directly perceived by pedestrians and may influence their walking experience significantly. Microlevel walkability is of a finer scale of measurement and is bound to heavily influence walking, as it is intrinsically ‘micro behavior’ (Park et al., 2014). Such a scale is composed of elements that can be improved in a short-medium time perspective (B. E. Saelens, Sallis, Black, et al., 2003) in cost-beneficial interventions (Knell et al., 2019). Micro-scale features are most commonly studied through the street segment unit of analysis using tools, i.e., audits, for in-person or online observational data collection (Blecic et al., 2020; T. Pikora, 2000).

Despite their relevance, such fine-grain observation methods are generally understood to be resource and time consuming (Blecic et al., 2020), are subject to bias (H. M. Badland, Opit, et al., 2010), inconsistent across different raters, and unfeasible or inefficient for extensive scale field observation (L. Yin & Wang, 2016a). Such onerous limitations are the most significant impediments for the micro-level usage in combination with macro-scaled characteristics, a long-foreseen recommendation (Maghelal & Capp, 2017).

Literature evidence indicates that omitting street-level characteristics from walkability studies might lead to over-estimations, for example, in areas where heterogeneity in sidewalk provision exists (Larranaga et al., 2018). In this sense, more research is needed to properly

incorporate microlevel built environment for the composition of multiscale tools, and more importantly, ascertain the benefit of including them (Arellana et al., 2019).

Advancements in technology might be highly beneficial if successfully applied to incorporating the micro-scale to macro-scale walkability analyses (Zhang et al., 2018). In recent years, online urban images from products like Google Street View (GSV) have become increasingly available, depicting the user's eye-level perspective (Middel et al., 2019). Through these tools, big data and machine learning approaches become possible for urban studies (Doersch et al., 2012; Moosavi, 2017; Shen et al., 2018). This type of procedure can be a more efficient alternative in measuring street-level features, as they make it possible to obtain a large data volume for analyses efficiently.

In this context, approaches using computer vision have gained space. Machine learning methods of obtaining data from the urban environment reduce the time spent on local surveys while allowing for a larger volume of data (Zhang et al., 2018). However, traditional machine learning is limited to hand-crafted features to describe a problem and its ability to deal with high dimensional data (e.g., pixels of an image obtained through GSV), resulting in the problem known as Curse of Dimensionality (Poggio et al., 2017). In contrast, deep learning methods present multiple layers of data processing and representation through the composition of non-linear modules that transform the initial representation (e.g., pixels of an image) into elements of a higher level of abstraction (Lecun et al., 2015).

Deep learning techniques, such as Convolutional Neural Networks (CNN), have advanced with computer vision applications in many research domains, including in the identification of compositional elements of the built environment such as aesthetics, building quality levels (Tan et al., 2017), urban greenery and canopy parameters (Liang et al., 2017). These applications have not yet been employed in comprehensive walkability studies and promise to be of great relevance to the research field.

Over more recent years, Deep Learning to measure walkability related urban features through semantic segmentation tasks (Koo, Guhathakurta, et al., 2022; Koo, Guhathakurt, et al., 2022) and classification tasks (Adams et al., 2022) have grown in popularity (H. Zhou et al., 2019). Deep learning models can learn to identify parts of an image according to classes (e.g., human, vehicle, construction, etc.), therefore, can be trained for streetscape analyses without any human image labeling involved. The automatic extraction of urban attributes that might compose influences on behavior, as an alternative for micro-scale walkability audits, is then made possible. These applications have not yet been employed in comprehensively

walkability studies in LMCI and promise to be of great relevance to the research field in this context, where there is a chronic lack of support to maintain costly assessments of walkability longitudinally.

The plethora of research on walkability indicates a relevant variance in walking across different populations (Dean et al., 2020). In this sense, another research gap yet to be filled concerns the assumption that walkability is an inherently contextual urban quality (Christian et al., 2011). State-of-the-art research indicates that walkability measurement methodologies might not be replicable across different scenarios. Instead, they might have to be developed for specific contexts (Shashank & Schuurman, 2019). Few studies have been conducted in the Global South, reflecting the lack of scientific scrutiny received by walkability in such context (Arellana et al., 2019). Further, especially in countries ranked between low and middle income, such as Brazil, active travel promotion is paramount due to the rapid expressive increase in population (Boutayeb & Boutayeb, 2005) and deep social disparities.

This relevance increases in urbanized areas other than big city centers since active travel in Brazil has shown to be more relevant as city sizes decrease (Associação Nacional de Transportes Públicos - Antp, 2018). According to the National Public Transport Association, the participation of Non-Motorized Transport (bicycles and on foot) in Brazil increases with the reduction of the municipality's size, being that those with 60 to 100 thousand inhabitants - considered to be medium-small cities- have higher percentages of active commuting. This points to the understanding that walking and walkability analyses are needed in medium and small cities mobility policy strategies.

Considering the topics presented above, there is a clear need for a greater understanding of multiscale walkability for medium-small Brazilian towns, attenuated by our pandemic-state, for tailored mobility policies. Therefore, this study aims to fill the research gap regarding built environment multiscale walkability in our unexplored context.

### 1.1 Thesis and research question

The statement to be tested —**the thesis**— sets forth that: *An efficient multiscale neighborhood-level walkability analysis for medium-small Brazilian Towns can be constructed through the use of deep learning techniques.*

Giving rise to a **research question**: *How can multiscale built environment attributes be incorporated in neighborhood-level walkability analyses for medium-small Brazilian Towns?*

## 1.2 Research Objectives

Leading to **the main objective**: *To elucidate on multi-scale walkability by designing neighborhood-level walkability models for medium-small Brazilian Towns using deep learning techniques.*

Under the light of this primary objective, **specific objectives** are required to capture adequate variables and suitably analyze multiple scales of walkability in the medium-small Brazilian town reality:

- 1) The first specific objective is to **Analyze** urban form to attend to local contexts and spatial nuances.
- 2) A second aim is to **Explore** deep learning and statistical methodologies for both data gathering and modeling.
- 3) A third aim is to **Assess** the benefits of including the micro-scale in macro-scale analyses, considering its relationships with walking.

The novel proposition of the research lies in the elucidations obtained in the process of designing the multiscale walkability models, and not necessarily in design artifacts themselves. The contributions this research will provide to the walkability field of study include (1) a better understanding of the walkability phenomenon in the Brazilian context; (2) the proposition of solid theory building through a novel method for gathering and analyzing walkability data and (3) the delivery of politically relevant evidence to guide urban planning strategies towards positively influencing active travel.

## 1.3 Research Outline and Thesis Structure

The approach to achieving the thesis and objectives was structured into a five-chapter process, each encapsulating a unique stage:

- **Introduction**: This chapter lays the groundwork for the research, providing an insightful contextualization of the research paradigm and presenting an overarching outline of the study.
- **Literature and Conceptual Review**: Serving as the second chapter, this part delves into a review of relevant literature and concepts, particularly focusing on urban form and

walkability, their various scales and methods, and the application of computer vision in urban studies.

- **Methodology and Data:** The third chapter explores the data and methods employed. It is strategically subdivided into three sections. Firstly, it contextualizes the case studies, walking level data, and the selected units of analysis. Next, it introduces the first phase of the methodological procedure regarding macroscale walkability. The second phase then examines microscale walkability, followed by the third phase which sheds light on the comprehensive Walkability modeling of the multiscale walkability index.
- **Discussion and Preliminary Results:** The penultimate chapter presents a discussion along with findings of the study.
- **Conclusion:** The final chapter provides final considerations, highlighting the strengths and contributions of the study, acknowledging limitations, and proposing potential avenues for future research.

As a methodological approach, considering this research as an exemplar of the ‘design sciences’ paradigm (Aken, 2004), being prescription-driven to solve problems faced in the real world (Lukka, 2003), Design Science Research (DSR) was selected. The investigation proposed here results from a multi-disciplinary and widely complex problem regarding urban studies, socioeconomic characteristics, and policy environment. The final outputs from the proposed DSR effort will be both a model (March & Smith, 1995), and more importantly, theory building, as the methodological construction of an artifact is an object of theorizing in itself (Vaishnavi & Jr., 2008).

Considering the phenomenon under investigation as real-life and contemporary, therefore indissociable from its contextuality, the adequate research strategy is the case study (R. K. Yin, 2001). Further, to deepen the understanding of relevant walkability measures, cross-validating findings in multiple empirical studies are necessary. Three different sized cities (approximately 60, 100, and 500 thousand/inhabitants) were selected: Rolândia, Cambé and Londrina, PR/Brazil. All of which have recently developed its mobility plans in 2017, 2017, 2020, respectively. Therefore, extensive secondary databases on walking trips were provided by the city administrations of the case studies.

For the consolidation of the research processes, a typical design science research effort is proposed: (1) *Awareness* of Problem, which results in a proposal for research such as what is being presented in this document; (2) *Suggestion*, which is a inventive step wherein new functionality is envisioned; (3) *Development*, basically consisting of the design development.

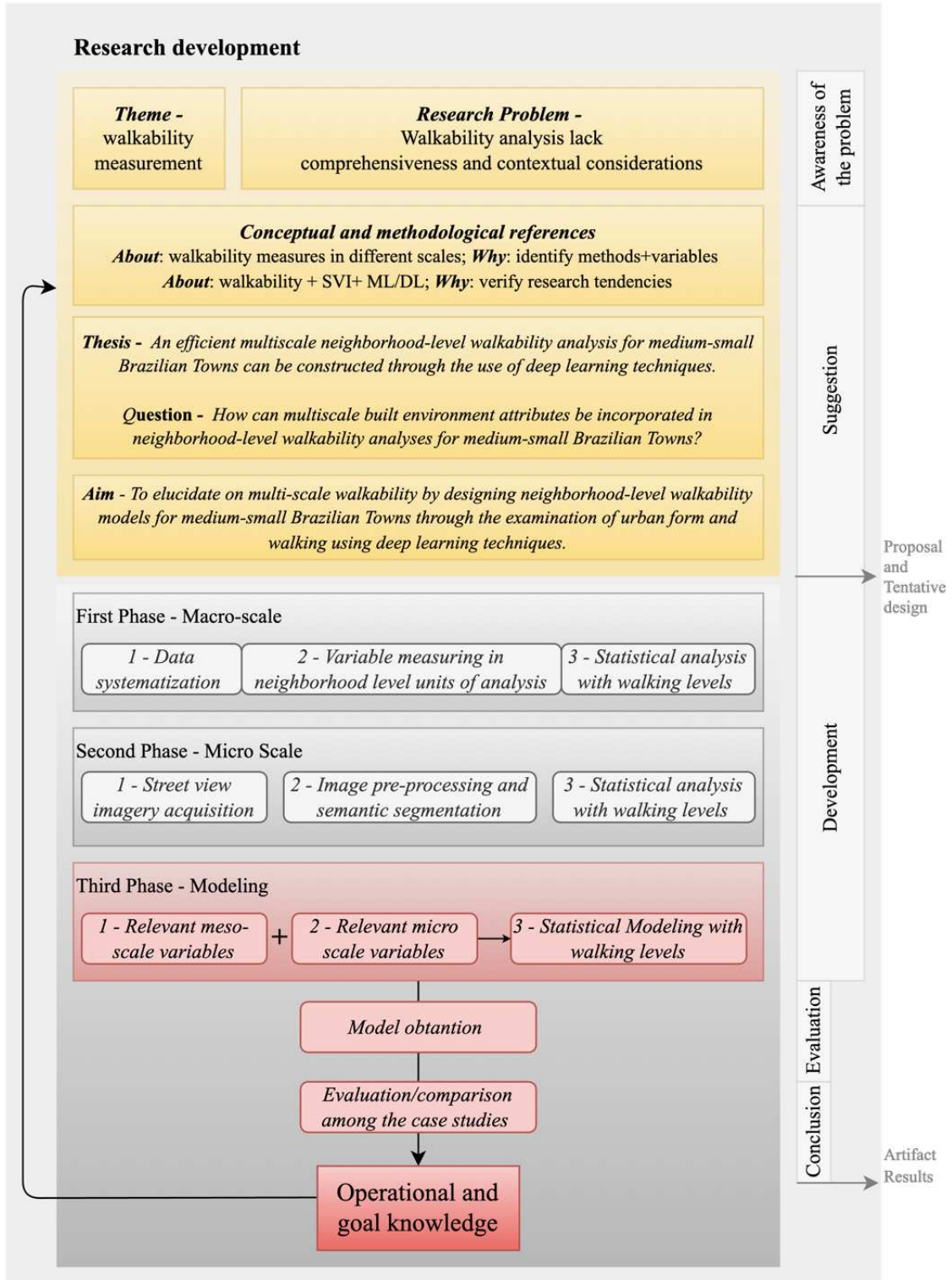
This step is further divided into three phases considering each walkability scale and comprehensive modeling; (4) *Evaluation*, where the artifact, in this case, the walkability models, are evaluated; (5) *Conclusion*, where research effort, if successful, is consolidated, at the end of the research. This process defines a flow of understanding of the research problem that possessed a practical-theoretical relevance. Also, the developing and evaluating of a possible solution to such a real-world problem, and finally reflecting upon on the outcomes of the research process (Vaishnavi & Jr., 2008). A research outline graphically explaining the development of the research is available in Figure 1.

The first phase of our methodological development (3) aims to examine the macroscale walkability variables incorporated in the traditional Walkability Index proposed by Frank et al., (2009) and the Space Syntax Walkability measure put forth by Koohsari et al., (2016b). We will focus on the following constructs within the built environment: Residential density, Retail floor area ratio, Intersection density, Land use mix, Space syntax metrics. The necessary data for these analyses have been either supplied by city administrations and systematized by our team within a GIS environment or collected by our researchers using the ArcGIS 10.6 software. The data aggregation process was conducted at both the zone and census tract neighborhood representation scales.

In a second phase, the methodology for quantifying walkability at the microscale level was selected through the systematic consideration of the specific scientific literature applying walkability analysis, street view imagery, and deep learning. From a sampling of street points 100m apart on the street network, GSV images will be collected from its API. Such images were processed using a Transformer based neural network model selected and trained for semantic segmentation in the Cityscapes dataset. Cityscapes (Cordts et al., 2016) is a large-scale dataset, which contains high-quality pixel-level annotations in 19 semantic labels from 7 categories (ground, construction, object, nature, sky, human, and vehicle) of 5000 images collected in street scenes from 50 different cities.

In the third phase, multi-scaled analyses of walkability were conducted using statistical modeling, with walking levels considered as a dependent variable and relevant micro and macro level variables identified in previous phases serving as explanatory variables. The results from this process, carried out across three different case studies, were then evaluated for their similarities. This methodological approach will enable us to achieve the proposed operational knowledge.

Figure 1 – Research Outline.



Source: Organized by the author, 2023

## 2 BIBLIOGRAPHIC REVIEW

Active behaviors such as walking are influenced by various factors, levels of determinants, and their interplay (Bauman et al., 2012). Individual variables are widely studied (SALLIS et al., 2016), yet environmental factors, despite their acknowledged impact on behavior, have received less research attention (Bauman et al., 2012). As the issues associated with motorized transportation escalate (Murray et al., 2014), the characteristics of urban environments have prompted researchers to explore the influence of urban form on travel behavior (Campoli, 2012). These studies have become increasingly central to discussions about sustainability and the heart of city planning strategies in higher income countries (Gilderbloom et al., 2015).

Walking not only contributes to health but is also fundamental to sustainable mobility. It reduces the reliance on motorized transportation, thereby diminishing environmental impacts. As a mode of transportation, walking is resource-efficient, inexpensive, quiet, and non-polluting (Gehl, 2013). However, it's crucial to recognize that, in low- and middle-income countries (LMICs) as well as certain high-income country settings and communities, active travel is often a necessity rather than a choice. In these contexts, a significant portion of physical activity, including walking, is not the result of voluntary decisions, but rather a direct consequence of economic necessity (Salvo et al., 2022).

While distinctions exist between walking out of choice and out of necessity are still being delineated, scholarly works often emphasize the differing influences on and behaviors between recreational and utilitarian walking. Both types play a role in promoting physical activity and ultimately, sustainability. When considering utilitarian walking, it typically has a purpose, beginning and ending at different places. On the other hand, recreational walking doesn't have a set destination and starts and finishes at the same point (B. Kang et al., 2017). Furthermore, the way the built environment affects walking can vary based on its purpose, as noted by various studies (Saelens and Handy, 2008; Sugiyama et al., 2012).

One of the strategies to evaluate the built environment is walkability, defined as “[...] the extent to which characteristics of the built environment may or may not be conducive to walking for either leisure, exercise or recreation, to access services, or to travel to work [...]” (Leslie, Coffee, et al., 2007, p. 91). There are two types of walkability in terms of the scale of measurement: micro and macro-level walkability (Park et al., 2015). Micro-level walkability consists of the walking environment that pedestrians directly perceive. Specific street-level characteristics comprise such walkability scale, such as the presence of trees, the shape of

buildings, their arrangement, the width of sidewalks, the quality of streets, and, fundamentally, sidewalk quality. These features are thought to have a more instantaneous influence on pedestrian perception and can be improved in a short and medium time perspective (B. E. Saelens, Sallis, Black, et al., 2003). Testing micro-level walkability as one of the determinants influencing walking travel behavior could be very important for policymakers as improving micro-level walkability has excellent potential to be a cost-beneficial intervention tool to promote walking.

The quality of the micro walking environment has long been of interest to urban designers and planners. Though the decades, urban designers and theorists have introduced various arguments that relate to street-level walkability. Some examples are Jacob's "eyes upon the street" (J. Jacobs, 1961), Alexander's degree of "enclosure" (Alexander et al., 1977), Appleyard's "livable streets" (Appleyard et al., 1981), and Jan Gehl's "soft edges" concept (Gehl, 1986, 1987). Currently, there is an effort to objectively measure and quantify such design attributes in a comprehensive way. Therefore, micro-scale walkability measurement tools have been developed (R. Ewing et al., 2006; T. Pikora, 2000; B. E. Saelens & Sallis, 2002; Sallis, 2016).

On the other hand, there are macro aspects of general urban form that may be relevant to walking and subsequent positive health outcomes. This walkability scale has been evaluated by measuring urban form attributes such as density, land use mix, and street patterns. (Cervero & Kockelman, 1997) propose in their seminal work the concept of the 3Ds: Density, Diversity, and Design. According to these authors, neighborhoods with high population density, diverse land uses, and pedestrian-oriented design are more likely to facilitate active travel choices, contributing significantly to overall physical activity. The concept of the 3Ds encompasses the following walkability components, often calculated as their sum: residential density (density), land-use mix (diversity), intersection density (design), and retail floor area ratio (design). Health behavioral ecological models concentrate on the extensive potential reach and long-term impacts of environmental factors, which rely on both macro structural characteristics such as street interconnectivity and land use mix (Brownson et al., 2009; Sallis, J. F., Owen, N., & Fisher, 2008) and microscale ones, which could affect the pedestrian experience (Cain et al., 2014; Sallis et al., 2011). Academically, it can be affirmed that a more considerable amount of studies on walkability have focused on either macro or micro-scale environmental variables separately (Kim et al., 2014). This can lead to falsely grounded evidence as both scales are considerably important in influencing walking behaviors and perception. Despite the

aspirations of creating healthy and sustainable cities in lower-income countries, there are greater limitations in policy data compared to those in higher-income cities. Furthermore, only a handful of these cities have quantifiable standards and targets to accomplish these goals (Giles-corti et al., 2022).

Considering such theoretical and methodological issues, the following chapters further discuss walkability in both scales and present a possible novel path for simultaneously and effectively incorporating them in the same analysis.

## **2.1 The macro scale approach to Walkability**

Walkability is considered a crucial element in supporting and improving the flow and mobility of pedestrians. In this sense, research on walkability has proliferated in the last decade, and with that, the emphasis on how to quantify walkability has also grown (H. Wang & Yang, 2019b). The most accepted approaches to quantify walkability are composite walkability indices (WI) that can calculate the combined effects of different aspects of the built environment (Arellana et al., 2019).

These composite indices, which sum and weigh variables, reflect various aspects of the built environment that enhance walkability. These indices were developed based on the research insight that individual variables are not as robust when used alone, compared to when combined. Thus, walkability emerges as a result of numerous characteristics, each acting as a predictor of health (Shashank & Schuurman, 2019). Indices are thought to capture the relationship of built environment characteristics while minimizing spatial collinearity and easing the communication of results (Brownson et al., 2009).

The WI serves as a community measure of walkability, assigning a value to neighborhood suitability to walk (Maghelal & Capp, 2017). By doing so, neighborhood WI allows for the comparison of local spatial units, variously referred to as community or residential environment for the investigation of the locus of social interactions and local behaviors (Macintyre et al., 2002; Sabel et al., 2013).

In 1997, the first composite walkability index was proposed by Cervero and Kockelman, (1997). However, the most widely accepted index in 2009, proposed by (L. D. Frank et al., 2009). This initiative was then followed by many well-known indices in walkability studies and some context-specific indices (Buck et al., 2011; Glazier et al., 2013; Stockton et al., 2016). A plethora of systematic reviews on walkability measuring - from multiple perspectives - have been conducted (Elshahat et al., 2020; Hilland et al., 2020; Maghelal & Capp, 2017). The

general conclusion is that indices are a clear literature tendency to evaluate walkability (Arellana et al., 2019).

Most walkability indices are conceived considering two fundamental factors understood to influence walking: *Connectivity* (directness of travel) and *Proximity* (distance) (Leslie, Coffee, et al., 2007). *Proximity* refers to the distance between the start of the journey (i.e., where one is) and the destination (i.e., where one is going). *Proximity* is mostly characterized by residential density, retail floor area ratio, and land-use mix. On the other hand, *Connectivity* characterizes the ease of travel between the origin (e.g., households) and destination (e.g., stores and employment). High connectivity implies that the route distance closely mirrors the straight-line distance, providing multiple alternative routes to the same destination within the street network. Conversely, low connectivity is characterized by large block sizes and obstacles that hinder direct travel (L. Frank, 2000; B. E. Saelens, Sallis, & Frank, 2003).

Residential Density, a key measure of proximity, serves as a proxy for shorter and more accessible walking trips. It has been consistently associated with positive outcomes for utilitarian walking (B. E. Saelens, Sallis, & Frank, 2003). By accommodating a sufficient number of residents, residential density can reduce the number of motorized trips and increase the number of walking trips (Cervero & Kockelman, 1997) thereby supporting the feasibility of urban services like local markets, schools, and public transportation. However, while the establishment of compact, high-density developments is advocated in contemporary urban planning, it often conflicts with various sociocultural contexts, particularly those in middle-income countries (Carmona et al., 2010).

Land use mix is perhaps the most recognized measure of proximity. It is associated with a higher presence of non-residential destinations in an area where diverse land uses are intertwined, leading to an increased likelihood of pedestrian movement (Cerin et al., 2007). Its conceptual roots can be traced back to the diversity construct proposed by Jane Jacobs (J. Jacobs, 1961), which posits that vibrant urban spaces result from a rich interplay between various uses. While diversity, and hence land use mix as a variable, is closely tied to the urban quality known as Vitality, it is within the realm of walkability research that the land use mix concept is most extensively applied (Frumkin et al., 2004). It is conceptualized primarily in quantitative methods when assessing functionally different uses' heterogeneity (Leslie, Coffee, et al., 2007). Currently, mixed land use is most widely quantified (L. D. Frank et al., 2004; Gebel et al., 2009; Grasser et al., 2016; S. M. Lee, 2010) as evenness, represented by the proportion of different land uses distributed in an area and quantified by variations of the

entropy equations proposed by Shannon (Shannon, 1948). Evenness of uses displays the proportion of uses' distribution, reflecting the degree of difference (Hajna et al., 2014)

The Retail Floor Area Ratio (Retail FAR), the ratio between the sum of commercial building floor area and the total commercial land area, is an indicator naturally related to land use mix (L. D. Frank et al., 2009). It was developed as a representation of more choices of destinations where products and services can be bought (Leslie, Coffee, et al., 2007). Further, as a pedestrian-oriented neighborhood design measure, it aims to identify the "pedestrian-unfriendly" architecture with large hostile parking lots where it is less likely for retail parcels to have a high retail floor area ratio (Cervero & Kockelman, 1997). This measure is closely related to major retail chains in North American cities and shopping malls (Leslie, Coffee, et al., 2007). It can be viewed as a derivation of the early parking ratio metrics, showing the relationship between the parking space and the space occupied by retail buildings (Gibbs, 2012).

*Connectivity* quantifies the relation between origin-destination and is argued to be an urban design measure that underpins a walkable neighborhood (Koohsari et al., 2016). Urban travel happens on streets, affecting travel patterns directly (Grasser, 2014). More direct routes, provided by linked street networks (Frank et al., 2009), are prerequisites for increasing pedestrian movement (Ellis et al., 2015). Several empirical findings indicate strong positive associations between walking, especially for transport, and street connectivity most expressed by intersections density (Berrigan et al., 2010; Oakes et al., 2007; Sugiyama et al., 2012). This metric is often weighted by a factor of two within common WI, such as in the prominent walkability study proposed by Frank et al. (2009). This environmental feature, however, is generally very distinct in Brazil (Leão, 2017; Leão & Urbano, 2020). Motomura (2017) also pointed out that in the Brazilian outskirts, locus of social housing, higher intersection density is present despite its general low walkability, as these areas are characterized by small lots and long-narrow blocks.

Compared to Intersection Density, integration - being a well-known *Connectivity* measure from the Space Syntax theory - considers how streets connect, becoming more or less accessible to each other. Space Syntax can be interpreted as a complex communication measure based on the spatial relationships within a network. (Hillier & Hanson, 1984) (1984), the theorists behind Space Syntax, argue that the movement of pedestrians is primarily dependent on the spatial structures generated by society. One of the proponents of the theory, (Hillier et al., 1993), indicated that the primary driver of pedestrian movement is the street configuration.

The relationship between the quantitative measures suggested by the theory has shown high correlations with the movement of pedestrians in many empirical studies, including in Brazil (Leão & Urbano, 2020; Zampieri, 2012).

Despite the considerable variation in how walkability is operationalized, the relationship between all these variables (i.e., residential density, land use mix, retail floor area ratio, intersection density, and integration) and walkability is widely recognized (Table 1). These are the primary attributes that transportation and urban planning researchers have investigated to understand how neighborhood macro characteristics relate to active travel.

However, these analyses, largely based on proximity to destinations and connectivity, capture the practicality of walking but not the aesthetic pleasure or the walker's experience. For example, if a restaurant is within walking distance but the route to it is dark and uninviting, will people be inclined to walk there? Drawing on these considerations, the potential incorporation of streetscape composition variables into the macro-level Walkability Index could yield more accurate indicators of walkability (C. W. Harvey, 2014; Liu et al., 2017a).

Table 1 - Environmental characteristics and relationships to walking behavior.

| Type         | Environmental component  | Implied Relationship with Walkability  | Necessary data                           |
|--------------|--------------------------|--|--|
| Proximity    | Residential density      | <ul style="list-style-type: none"> <li>• Density Improves accessibility to complementary uses</li> <li>• Associated with increases in retail and service variety,</li> <li>• Results in shorter distances between destinations</li> </ul>  | Residential location data                |
|              | Land Use                 | <ul style="list-style-type: none"> <li>• People who live near multiple and diverse retail opportunities tend to make more frequent, more specialized, and shorter shopping trips, many by walking</li> <li>• People who live farther away from retail are more likely motorized transportation for every-day purchases</li> <li>• land use mix generates more interesting built- form increasing esthetic value making routes more conducive to walking</li> </ul> | Land use data                            |
|              | Retail floor area ratio  | <ul style="list-style-type: none"> <li>• More options for destinations where goods and services may be purchased</li> <li>• More local employment opportunities can be reached by walking</li> </ul>   | Retail location and building projections |
| Connectivity | Intersection-density     | <ul style="list-style-type: none"> <li>• Higher intersection densities provide people with a greater variety of potential routes</li> <li>• Higher connectivity provides easier access to major roads where public transport is available</li> <li>• Shorter times to get to destinations</li> </ul>   | Street centerline and intersections data |
|              | Space Syntax Integration | <ul style="list-style-type: none"> <li>• More integrated areas have the potential for the development of commercial activities</li> <li>• More integrated areas support of land use mix through pedestrian movement</li> </ul>   | Axial lines from street centerline data  |

Source: Organized by the author, 2020. Based on (Leslie, Cerin, et al., 2007), Saelens et al. (2003), Leão (2019), and (Koohsari et al., 2016).

## 2.2 The micro scale approach to Walkability

In the process of defining walkability operationally, a large number of studies have limited their environmental evaluations to macro-level data. Despite the considerable progress in developing objective methodologies for measuring walkability at the macro level, micro-features - the detailed aspects of urban spaces - have a substantial impact on their walkability. Taking into account the speed and character of walking, pedestrians are generally seen as more attentive and responsive to environmental characteristics than those traveling in motor vehicles. Consequently, the structural components of streets are thought to significantly shape pedestrian behavior patterns (Clifton et al., 2007; Talen, 2002). Alternatively, when viewed from a perspective that considers spatial inequities, these structural elements are crucial in creating either high-quality or substandard environments for those who rely on walking out of necessity.

Further, experimental studies have been conducted stating the relevance of street composition on walking behavior (Knell et al., 2019). From a more practical perspective, the street-level composition is much more easily modified and has had the cost-effectiveness of interventions recognized (Gunn et al., 2014; Laine & Kuvaja-ko, 2014). Over the last two decades, urban planners and health researchers have created an array of methods for assessing the physical conditions of neighborhoods at street level through audits (Cain et al., 2018; Millstein et al., 2013; T. J. Pikora et al., 2002). These measures deploy systematic observations to objectively evaluate features of the urban environment at the street level.

Typically, audit tools are used to assess physical attributes best understood through direct observation, such as aesthetics or maintenance, along with a broader objective quantification of features. These tools generally necessitate in-person inspections conducted by researchers trained as evaluators to gather data. A standardized methodology for systematically coding features during a walking or driving assessment, complete with definitions, is put in place (Brownson et al., 2009). The most frequently used unit of analysis is the street segment, which is defined as a block face situated between two intersections. Street segments are typically sampled, as auditing entire neighborhoods, barring a few exceptions, is impractical. However, segment-level observations offer an advantage over macro-level analyses as they refer to more detailed characteristics (S. Lee & Talen, 2014).

From the seminal reviews proposed (Brownson et al., 2009; Moudon & Lee, 2003)), to more recent work (Aghaabbasi et al., 2017; S. Lee & Talen, 2014)), walkability audit methods have been well documented with their evidence consistently and systematically registered. Several studies have made the argument that audits are capable of assessing the ‘unmeasurable’,

such as the imageability or the legibility of an area (R. Ewing & Handy, 2009). Some have examined how perceptions of street-level features coincide with their objective calculations (Gebel et al., 2011; McCormack et al., 2008). Further, some, have focused on walking categories, such as leisure or transportation (C. Lee & Moudon, 2006).

Additionally, numerous audit-based instruments have been developed (Boarnet et al., 2006; Brownson et al., 2004; Cain et al., 2018; Clifton et al., 2007; Emery et al., 2003). The Systematic Pedestrian and Cycling Environmental Scan (SPACES) (T. J. Pikora et al., 2002); the Pedestrian Environmental Data Scan (PEDS) (Clifton et al., 2007), and MAPS (Cain et al., 2018); are some of the most widely applied of instruments (S. Lee & Talen, 2014). As originally articulated by Moudon and Lee (2003), most audits primarily concentrate on 'spatio-physical' elements, such as the presence of sidewalks or trees. However, 'spatio-behavioral' aspects (e.g., pedestrian volumes and safety) and 'spatio-psychosocial' factors (e.g., perceived comfort, attractiveness) are often incorporated as well. These aspects are significant as they relate to the level of human engagement within the environment and people's subjective responses.

In this sense, most neighborhood audits contemplate physical measures of land use (e.g., presence of residences and retail) and sidewalks (e.g., presence of sidewalks). Also, behavioral linked elements such as streets and traffic (e.g., traffic volume); public amenities or supports for walking (e.g., presence of street furniture). Additionally, elements that deeply influence the individual perception of the urban landscape are eventually included, such as architectural characteristics (e.g., building height or enclosure) and maintenance or aesthetics (e.g., presence of litter or presence of graffiti) (Table 2) (Brownson et al., 2009).

However, as Brownson and colleagues (2009) put it: there are skills and trade-offs associated with using observational audit measures. One of the greatest difficulties of this approach is the time and resource intensiveness, leading to significantly narrow research areas (Hanibuchi, 2019; Rundle et al., 2011). The on-site approach is argued to be time-consuming, expensive, and burdensome (Moniruzzaman & Paez, 2012).

Table 2 - Summary of most common environmental features in micro-scale audits.

| <i>Aspects of Behavior-Environment</i> | <i>Categories</i>                      | <i>Items</i>   |
|--|--|--|
| <i>Spatiophysical</i>                  | <i>Land uses</i>                       | Land uses (types, intensities, destinations)             |
|  | <i>Walking path/sidewalks</i>          | Sidewalk presence  |
|  |  | Sidewalk qualities (materials, obstructions, uniformity) |
|  |  | Slope  |
| <i>Spatiobehavioral</i>                | <i>Vehicle-pedestrian interactions</i> | Natural barriers (ditch, creek)                          |
|  |  | Street supports for walking (crosswalks, traffic lights) |
|  |  | Traffic volume   |
|  |  | Parking (on and off-street)                              |
|  |  | Speed limits   |
|  |  | Segment/road connectivity                                |
|  |  | Road conditions (materials, uniformity)                  |
| Traffic calming (chokers, chicanes)    |  |  |
| <i>Spatiopsychosocial</i>              | <i>Safety and appeal</i>               | Lighting   |
|  |  | View/surveillance  |
|  |  | Aesthetics (incivilities, gardening, appeal)             |
|  |  | Unique markers/memorability                              |
|  |  | Architectural variety                                    |
|  |  | Enclosure  |
|  |  | Tree presence  |

Source: Organized by the author, 2020. Based On (Clifton et al., 2007; S. Lee & Talen, 2014; Moudon & Lee, 2003).

As an alternative solution to these problems, several studies have proposed virtual audits using services like Google Street View (GSV) that provide a virtual representation of the street environment for the conduction of “remote” and easier to execute audits (Rzotkiewicz et al., 2018a). These have been repeatedly tested as to their reliability, in most cases presenting strong reliability (H. M. Badland, Duncan, et al., 2010; Clarke et al., 2010; Pliakas et al., 2017).

Street view imagery does offer a clear option to assess the built environments for less time and cost (Rzotkiewicz et al., 2018b). However, it is still a manual human labor-intensive approach. It also still lacks in sample sizes due to the limit of human processing and still might be inefficient for larger-scale observations. Further, despite advances, virtual audits are still

possibly subject to bias. On a more specific note, image resolution can limit the identification of smaller objects to the human eye, possibly weakening the reached evidence. These arguments suggest that virtual auditing might not be the best alternative in all cases (Fry et al., 2020)

Despite numerous attempts to create reliable and cost-effective tools for collecting and systematizing street-level data, challenges related to limited geographical coverage and potential biases persist. The restricted geographical scope of street segments as units of analysis impacts the generalizability of findings, given the variability of physical features across different urban regions (Rzotkiewicz et al., 2018b). Additionally, research sites are traditionally selected through sampling, which could potentially lead to inaccurate conclusions considering the complex modeling of socio-economic correlates at the neighborhood level within street segments (Zhou et al., 2019b). Finally, regardless of the methodology used, audits have been criticized for often yielding inconsistent and unreliable measurements due to subjective interpretation, even when expensive rater training is implemented (L. Yin & Wang, 2016a).

### **2.3 Machine learning applied to urban analyses**

In recent years, online urban imagery from resources like Google Street View (GSV) has become increasingly accessible, providing a user's eye-level perspective (Middel et al., 2019). Some research highlights potential drawbacks in using such commercial imagery for walkability studies, noting that the images are captured from a vantage point higher than the typical pedestrian view (Biljecki & Ito, 2021a; Steinmetz-wood et al., 2019). However, this data remains of high value due to its numerous beneficial aspects. By leveraging these tools, we can incorporate big data and machine learning approaches in studies focused on the visual characterization of typological elements, use of high-resolution images for urban morphology research, and the measurement of urban qualities that possibly influence street scale walking behavior and well-being (Doersch et al., 2012; Liu et al., 2017a; Moosavi, 2017; Shen et al., 2018; Yin and Wang, 2016b).

When it comes to walkability, this type of procedure can be a more efficient alternative when compared to field surveys or audits (Ben-joseph et al., 2015), which in addition to possibly being prone to subjectivity, are more costly and not always safe for the researcher. (H. M. Badland, Opit, et al., 2010). Further, urban image acquisition methods such as GSV make it possible to obtain a large data volume for analyses (F. Zhang et al., 2018). In this scenario, works leveraging different urban applications through the automatic processing using deep learning of street view images have grown in the recent past, due to services from companies

such as Google, Baidu, and Tencent that provide high-resolution, geocoded, streetscape images worldwide (Rzotkiewicz et al., 2018a).

In this promising field, the initial studies focused on street-level urban greenery, canopy parameters, building quality or aesthetics, and general urban qualities, such as openness and enclosure. The most recent analyses focus on urban qualities understood as more ‘traditional,’ mostly theoretically based on Ewing and colleagues' seminal work (R. Ewing et al., 2006; R. Ewing & Handy, 2009). Such literature has heavily influenced recent studies on the quantification of built environment elements, where frameworks consider classic urban design qualities such as imageability, enclosure, human scale, transparency, and complexity, quantitatively as proxies for walkability. However, it must be emphasized that this is not a common foundation in the specialized walkability body of literature, despite this shift being considered a great “success” in walkability studies as these characteristics are beneficial for the improvement of measures (H. Wang & Yang, 2019b).

Deep learning methods present multiple layers of data processing and representation through the composition of non-linear modules that transform the initial representation (e.g., pixels of an image) into elements of a higher level of abstraction (Lecun et al., 2015). These techniques have advanced with computer vision applications in many research domains, including urban analyses. One of the significant scientific developments and research focuses on deep learning applications in urban studies has been street-level greenery, a widely recognized influence in behavior and well-being (X. Li et al., 2015).

Many studies aim to identify greenery on GSV images as a proxy for urban quality. In a first effort, (X. Li et al., 2015) verified that street-level urban greenery could be effectively analyzed through street view imagery (SVI). The same research trend (Long & Liu, 2017a) managed to calculate a previous green view index modification from over one million Tencent Street View pictures using semantic pixel-wise segmentation. Further, (Seiferling et al., 2017) applied image segmentation, and classification techniques on GSV images for quantifying the tree canopy cover at the street-level, and (Y. Zhang & Dong, 2018) quantified the street-visible greenness of residential neighborhoods through a *SegNet* trained for pixel-level semantic segmentation.

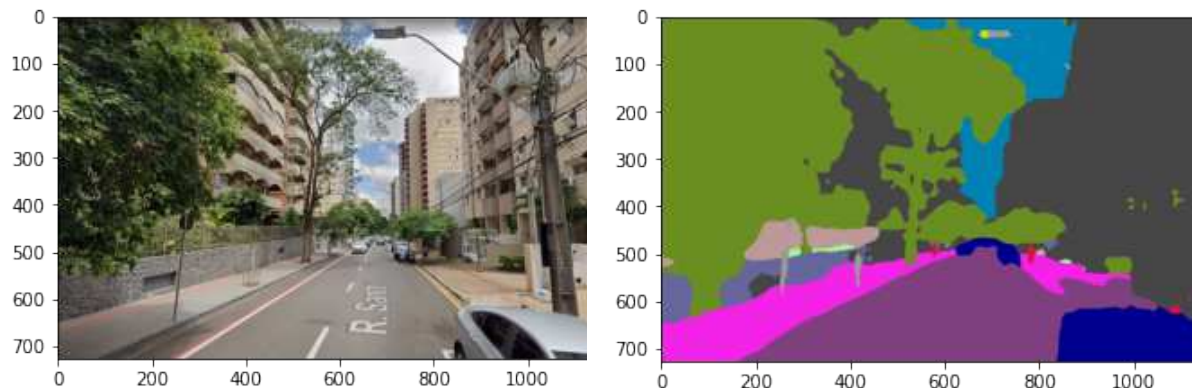
Canopy parameters have also been effectively estimated with street view images in the recent past. Studies such as those proposed by (Zeng et al., 2018) and (Liang et al., 2017) aim to automatically identify the amount of sky of street view imagery as a parameter for radiation and thermal environmental assessment and of the human visual perception of urban landscapes.

Building quality and aesthetics from urban landscapes are also relevant topics being addressed in current research. The work proposed by Lu et al. (2017), to automatically evaluate the quality of the urban environment, measured maintenance quality of building facades and the continuity of street walls. Whereas in a more aesthetic-related approach, (Novack et al., 2020) detected building facades with graffiti artwork from the perspective that it is a cultural asset and an essential aspect of a city's aesthetics. The studies on the recognition of architectural identity or style are mention-worthy. From the seminal work 'What makes Paris look like Paris' from Doersch et al. (Doersch et al., 2012), many studies have developed such ideas (Marijana, 2020; Obeso & Benois-pineau, 2019; Shalunts et al., 2011; Xu et al., 2014), but are not the focus of this research despite such aesthetic attributes being eventually discussed in walkability measurements (Saelens & Sallis, 2002).

Despite this, only a handful of studies have explicitly employed deep learning or machine learning methods in the analysis of walkability. As of 2022, no research had thoroughly addressed the systematization of walkability based on quantification through deep learning in LMICs for entire cities. Prior to this, a few studies implicitly addressed the issue by considering overarching characteristics of the built environment that constitute walkability as a concept. Nevertheless, the pioneering efforts by Adams et al. (2022) and Koo et al. (2022a, 2022b) have advanced the direction of microscale walkability quantification.

From the initial success of Deep learning's Convolutional Neural Networks (CNNs) for image classification, beginning with the ImageNet (Krizhevsky et al., 2012), CNNs have become the dominant approach for almost all recognition and detection tasks (Lecun et al., 2015). Over more recent years, such breakthroughs of Deep Learning in image classification have been transferred to semantic segmentation tasks (L. C. Chen et al., 2018) and a shift has been observed with the rise of transformer models (Dosovitskiy et al., 2020). Prowess across a myriad of tasks, including semantic segmentation, has been demonstrated by these transformer models, challenging the dominance of conventional techniques, and potentially rendering them obsolete. Semantic segmentation aims to assign semantic labels to every pixel in an image (L. C. Chen et al., 2018), where percentages of segments for each type can be calculated (Figure 2). The models learn to identify parts of an image according to classes (e.g., human, vehicle, construction, etc.). Therefore, models developed for semantic image segmentation can be trained on streetscape datasets. The extraction of urban attributes that might compose influences on behavior, in an alternative for micro-scale walkability audits, could then be extracted.

Figure 2 - Image Processing Example Based on Semantic Segmentation.



Source: Developed by the author (2020). Based on Google Street View, (2020).

### 2.3.1 Deep learning using Street View Images for Walkability: state of art

In order to map this research field a systematic search of the literature, a method for gathering and synthesizing research results on a limited theme in a systematic fashion, was conducted (Ferenhof & Fernandes, 2016). The formal objective of the review, its explicit statement, was to *'verify the growth and countries of research on walkability or walking behaviors that analyze street view imagery through machine or deep learning'*. The aim is to ensure a comprehensive representation of the research field, minimize bias, and delineate any gaps in the evidence base. Considering such a non-formal character, this review has not been registered.

The eligibility criteria for the inclusion of a paper in the analysis were: (a) Published articles in which any walking behavior (for transport or leisure) in a sample of randomly selected individuals of any age, gender, or social group was analyzed or Published articles in which walkability is explicitly analyzed (in any of its definitions, scales or assessment methodologies) in a quantitative manner; (b) studies where machine learning or deep learning was used to compose walkability or walking behavior analysis; (c) studies must include the usage of street view imagery independent of source; (d) studies must apply machine learning or deep learning technique for the analysis of street view imagery. Journal or proceeding articles published in English without temporal restriction were eligible.

Initially, a research protocol was defined to guide the inclusion of quantitative studies. For such, the following English language descriptors were used to compose a search strategy carried out on 6<sup>th</sup> of August, 2023: `((walk*) AND ("machine learning" OR "deep learning" OR "image* segmentation" OR "computer vision")) AND ("street view" OR "street view image*")`. The first term, considering the asterisk in the root word "walk" aimed to identify studies

referencing walk, walkability, walking, walkable and further variations. The second search term links the possible artificial intelligence techniques, aiming to broaden options, including general terms and not specific models. The last terms filter studies that must necessarily utilize some sort of street imagery. The Boolean operators "OR" and "AND" were employed to combine words within and between the term's groups, respectively. The selection of articles was carried out without a temporal restriction or research area. The Information sources were the databases Scopus, Web of Science®, PubMed®.

The organization of the bibliographic portfolio, which allows for the consolidation and comparing of data, was conducted through software that organizes bibliographies and references (Mendeley). The standardization of the article selection was firstly based on exclusion by duplicity. After that the reading of all titles and abstracts was conducted.

In an initial search conducted in 2020 the study selection process identified a total of 21 papers were retrieved in electronic searches, in the year 2022 the same exact search was conducted and in the same process 54 were retrieved. Now, in the year of 2023 the literature search identified a total of 160 records across three databases: Web of Science (89), Scopus (65), and PubMed (6). After removing duplicates, we were left with 100 unique records.

It is noteworthy to mention that from the complete 100 records identified, it was possible to verify the expansion and diversification of this field of study. There has been a growing interest in topics such as urban vitality, quantification of green spaces and their impacts on perception and comfort, bikeability studies, and running environments. Further, notable contributions to this subject include systematic reviews around the usage of street view imagery for built environment assessments by Li, Y., Peng, L., Wu, C., & Zhang, J. (2022), Cinnamon, J., & Jahiu, L. (2021), and Biljecki, F., & Ito, K. (2021).

The exploration street view imagery for urban analytics and geospatial data collection reviews reveals a growing importance. One review highlights walkability's multifaceted nature, indicating that imagery provides crucial data for understanding and enhancing urban walkability (Biljecki & Ito, 2021). Similarly, another review underlines how street view imagery can measure walkability attributes of the built environment, including the quality of public spaces and neighborhood street walkability through indicators such as street greenery (Y. Li et al., 2022). The usage of panoramic street-level imagery in evaluating walkability and its impact on health and physical activity is also discussed (Cinnamon & Jahiu, 2021). Overall, this burgeoning application of street view imagery in assessing walkability is emphasized in these reviews and it is clear that they could shape urban planning and development in the future.

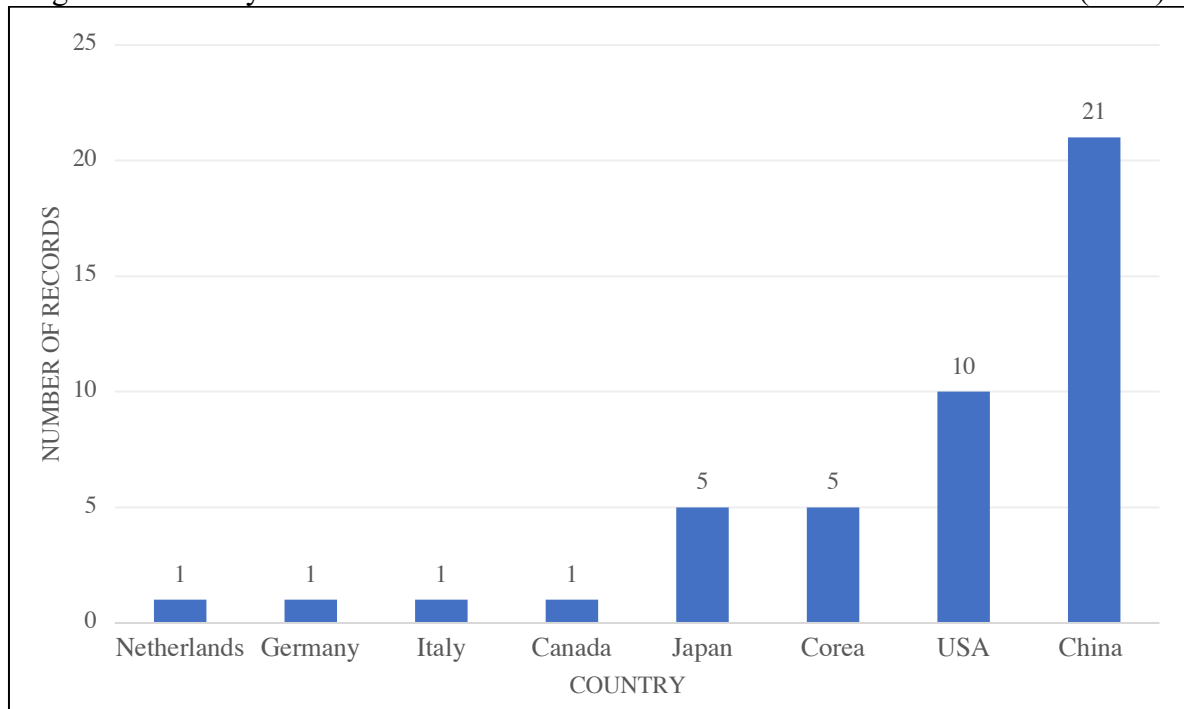
Following the filtering of titles and abstracts, **46 records** were identified as relevant in being specifically focused of walking or walkability, utilizing street view data and applying deep learning techniques. The geographical distribution of these studies was as follows<sup>1</sup>: High-income economies contributed 52.17% or 24 papers, while upper-middle-income economies accounted for 45.65% or 21 papers. Interestingly, one paper did not disclose its location. Figure 3 provides a more detailed representation of the country-wise distribution of these records, indicating that China and the USA are the primary contributors in this field.

There were no papers identified from LMICs focusing on walkability or walking behaviors with the application of street view images and deep learning (Figure 3). This only reiterates the great importance of advancing walkability studies in the global south, which significantly lacks evidence on walkability and walking behaviors. Such focus is paramount considering the urban dynamics, security issues, sidewalk invasion problems, and inadequate planning, a very different scenario from the dominant walkability research areas, making it challenging to improve the regional transferability metrics (Arellana et al., 2019). Of particular interest is a study conducted in Sao Paulo, Brazil by (Lucchesi et al., 2023), which utilized the discussed methodologies to analyze walkability. This study, however, was not identified in our systematic literature search. Despite this, its contributions represent a significant advancement in the field, especially within Brazil. It should be noted that the study's focus is primarily on walkability perception within a single neighborhood in the city. This emphasis suggests the potential for future research to expand in scope, considering a broader range of urban contexts.

---

<sup>1</sup> One of the 46 studies did not disclose the country or area of the case study conducted, therefore when accounting for location and income level of the countries of study, the sample size is of n=45.

Figure 3– Country distribution of research records obtained from literature search (n=45).

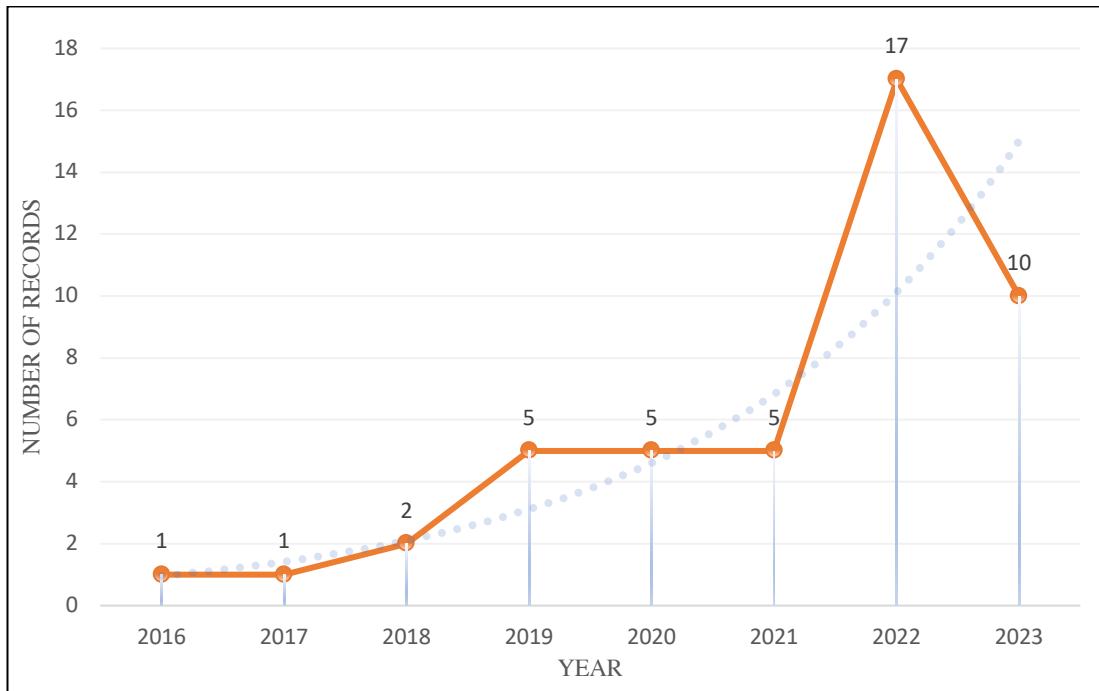


Source: Organized by the author, 2023.

A noteworthy observation from our review is the exponential increase in the number of publications in recent years, particularly in 2022 with 17 papers. In contrast, only 5 papers were published in 2021. In the ongoing year of 2023, there have already been 10 published papers, as shown in Figure 4. This significant increase suggests a rapidly growing interest in this area. As the field moves forward, the volume and depth of this research to continue expanding is expected, further enriching our understanding and application of deep learning and street view images on walkability analysis.

GSV is the predominant source of street view imagery. However, in Asian-centric studies, Baidu Maps Total View and Tencent Street View are frequently employed, given their extensive coverage of cities within China, therefore play a significant role in the domain of street view research, particularly within the Asian context (Cinnamon & Jahiu, 2021). All these services offer similar products and have been chosen considering their availability to researchers.

Figure 4– Yearly distribution of research records obtained from literature search (n=46).



Source: Organized by the author, 2023.

Street view images, typically drawn from the road network, are sampled using varying methods as per the research requirements, study area, and available street view data (Y. Li et al., 2022). Some studies have employed systematic sampling methods to determine street view locations, while others have adopted random or purposive sampling strategies (Biljecki & Ito, 2021). In terms of image retrieval, several studies have opted to combine four separate images to form a panoramic streetscape view. Alternative approaches include constructing a semi-panoramic view with three images or a smoother panorama using six images. The selection of geotagged points for image collection often occurred at intervals of 50 or 100 meters along the street network. However, a few studies chose points or intersections randomly. The quantity of sampled images varies drastically among studies, ranging from a few thousands (L. Yin & Wang, 2016c) to millions (Nguyen et al., 2020), demonstrating the diversity in methodological approaches within this research area.

There seems not to exist a trend for models, as they have varied dramatically among studies, it can be conjectured that there is a tendency to follow the most up to date and efficient pretrained model, which changes at a fast pace in the computer vision field. It seems that there is no silver bullet for selecting the right model, where they vary with the swift developments over the years. Initial studies applied variations of the VGG model (Simonyan & Zisserman, 2015; Yue et al., 2022), FCN8s (Shelhamer et al., 2017), DeepLab (L. C. Chen et al., 2018;

Lucchesi et al., 2023), and Pyramid scene parsing network (PSPNet) (Doiron et al., 2022; Koo, Guhathakurt, et al., 2022; Zhao et al., 2017).

Most of these works were based on convolutional neural networks, a more traditional architecture that mastered computer vision tasks until recently. Since the proposal of the Transformer architecture for machine translation (Vaswani et al., 2017), Transformer-based models has become the state of the art in many natural language processing tasks (Devlin et al., 2019; Raffel et al., 2019). Transformer models were first introduced in computer vision in (Dosovitskiy et al., 2020), where they proposed Visual Transformer, adapting the original design of the network for image inputs leading to impressive results on standard benchmarks for image classification. Following up, (Zheng et al., 2021) and (W. Wang et al., 2021) extended Visual Transformers to overcome its limitations and demonstrated considerable improvements over other methods on object detection and semantic segmentation.

When it comes to the training dataset adopted, the Cityscapes dataset (Lucchesi et al., 2023), the ADE20k dataset (B. Zhou et al., 2019), and Places365 dataset (Y. Kang et al., 2023) are common for pretraining a model combined with training on a created dataset for the characteristic of interest (B. Zhou et al., 2018).

Features of the built environment extracted from street view images often vary, reflecting the specific objectives of each study. Prominent features typically include aspects of the sky (indicating enclosure) and greenery. Some studies concentrate on a distinct class of features, whereas others take into account all classes present in the training data, frequently using this information to create theoretically informed indicators for particular attributes. However, a notable limitation arises when certain urban characteristics, such as greenery or sky proportion, are employed as standalone metrics for walkability. This reductionist approach may result in an absence of rigorous evidence due to the oversimplification of the multifaceted concept of walkability (H. Zhou et al., 2019).

Initial studies in the field utilized the existence of crosswalks and sidewalks as proxies for walkability (Nguyen et al., 2020), or composed a basic infrastructure composite encompassing pavements, streetscape components (such as greenery and enclosure), and social factors like crowdedness (H. Zhou et al., 2019). More recently, research has evolved to conduct in-depth, domain-specific analyses that, in some cases, aim to equate to specific pre-existing and traditional walkability measurements such as MAPS (Koo, Guhathakurta, et al., 2022).

It is important to highlight that recent research, including the work by Adams et al. (2022) and Ito and Biljecki, (2021) work on walkability and bikeability respectively, have

employed machine learning for automated analysis using classification tasks and a binary approach to semantic segmentation for specific attributes such as streetlights, bike lanes, and other street amenities. The fundamental premise behind this approach is that measuring the existence or non-existence of certain elements may be more significant than calculating their size, as pixel count might not convey the needed information. This approach aligns with traditional audits like MAPS (Ito & Biljecki, 2021), focusing on mimicking these audits rather than quantifying the presence of streetscape amenities. This approach has solid theoretical foundations and might be a relevant field of study to be explored in the LMICs context.

### 3 METHODOLOGY AND DATA

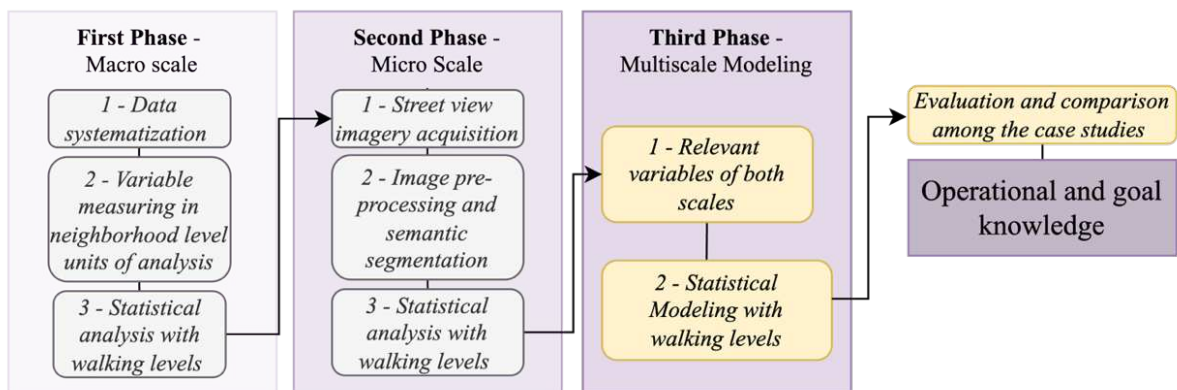
This research aims: *To elucidate on multi-scale walkability by designing neighborhood-level walkability models for medium-small Brazilian Towns using deep learning techniques.* Considering the phenomenon under investigation as real-life and contemporary, complex, and indissociable from its contextuality, the adequate research strategy is the case study (R. K. Yin, 2001). Further, to achieve a deep understanding of the phenomenon and generate more robust and reliable evidence (Gustafsson, 2017), cross-validating findings in multiple case studies were necessary. In this sense, three different sized cities (with approximately 60, 100, and 500 thousand inhabitants) were selected: Rolândia, Cambé, and Londrina, Paraná- Brasil. Through this approach, micro and macro level walkability characteristics were tested in multiple scaled cases.

Design Science Research (DSR) was adopted as a methodological approach, considering this study as fitting in the 'design sciences' category (Aken, 2004). The study proposed here is the product of a multidisciplinary and highly nuanced issue regarding urban studies and, further, is prescription-driven to solve problems faced in the real world (Lukka, 2003). The final outputs from the proposed DSR effort will be mainly theory-building stemming from the construction of models (March & Smith, 1995), from the understanding that the methodological construction of an artifact is an object of theorizing in itself (Vaishnavi & Jr., 2008). A typical design science research effort is proposed, composed of Awareness of Problem and Suggestion, presented in the first two chapters of this document. Design development, Evaluation and Conclusion, the final stages of the research, are in the following final chapters.

The methodological process development is divided into three phases (Figure 5), considering each walkability scale and a final step for general modeling. Initially, the

Macroscale analysis was conducted for each of the cities. Residential Density, Land Use Mix, Retail FAR, Intersection Density, and Space Syntax metrics were systematized, aggregated at the census tract/zone level, and statistically analyzed in their ability to predict walking levels according to origin and destination data surveyed for the urban mobility plans of each case study. Secondly, the microscale features of the built environment were analyzed. Image acquisition for semantic segmentation was obtained from GSV. These images were preprocessed and fed to a deep learning model trained for semantic segmentation. The extracted features were systematized, aggregated at the census tract/zone level and analyzed statistically in their ability to predict walking levels. Lastly, relevant features from the two previous phases, both macro and micro, for each city, were conjunctly statistically modeled with walking as a binary response to obtain three models that elucidate on the interaction between the different scale variables. Finally, the operational and goal knowledge was reached through the methodological construction of a multiscale neighborhood-level walkability models for medium-small Brazilian Towns through deep learning techniques.

Figure 5 - Development of the research process in three phases.



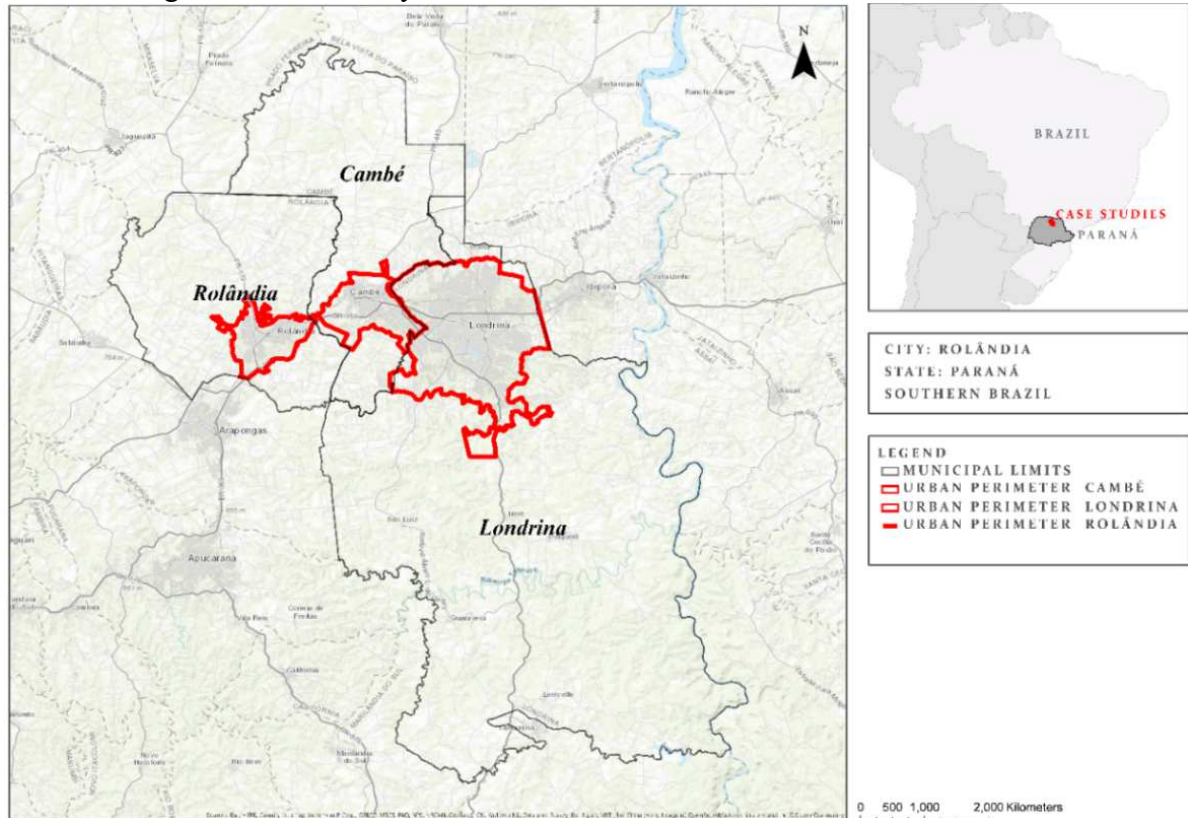
Source: Organized by the author, 2020.

### 3.1 Case studies: Walking data and aggregation.

The three selected cities are located in the northern region of Paraná, all being a part of the Londrina Metropolitan region (Figure 6). Rolândia, Cambé, and Londrina are cities founded by the *Companhia de Terra Norte do Paraná*, later *Companhia Melhoramentos Norte do Paraná*. They are settlements that follow the territorial planning for the undertaking and sale of plots destined, above all, to the coffee crops. The formation of a network of cities was a fundamental part of the colonization of this region. The established plan considered an average distance of 100 km between “big cities” – one of them being Londrina - that would serve as nuclei. Among them, smaller-sized cities were founded, with a maximum distance of 15

kilometers from each other, to establish commercial centers and intermediate suppliers – Cambé and Rolândia being some of these (Rego & Meneguetti, 2006).

Figure 6 – Case study location, Rolândia, Cambé and Londrina -Brazil.



Source: Organized by the author, 2020.

Rolândia was selected as a case study due to its representation of an average-sized town and data availability. The municipality of Rolândia in Paraná-Brazil has an extension of 454,174 km<sup>2</sup> and an estimated population in 2017 of 64,726 inhabitants (IBGE, 2018). Due to the lack of administrative infrastructure, georeferenced spatial data was not available. Therefore, the Georeferencing of data for the city of Rolândia-PR was conducted by the researchers. Walking data was secondary, extracted from the Rolândia Urban Mobility Plan (conducted in 2017 and 2018 under the responsibility of ITEDES. A total of 756 valid questionnaires was applied, representing 3.76% of the 20.065 Permanent Private homes (IPARDES, 2018). Two thousand seven hundred thirty-one (2731) trips were registered in the OD survey conducted in Rolândia. Due to raw data inconsistencies such as the misspelling of street names and missing information, a total of 2152 trips by any mode of transport were accounted for, being 394 of these trips being ‘by foot’.

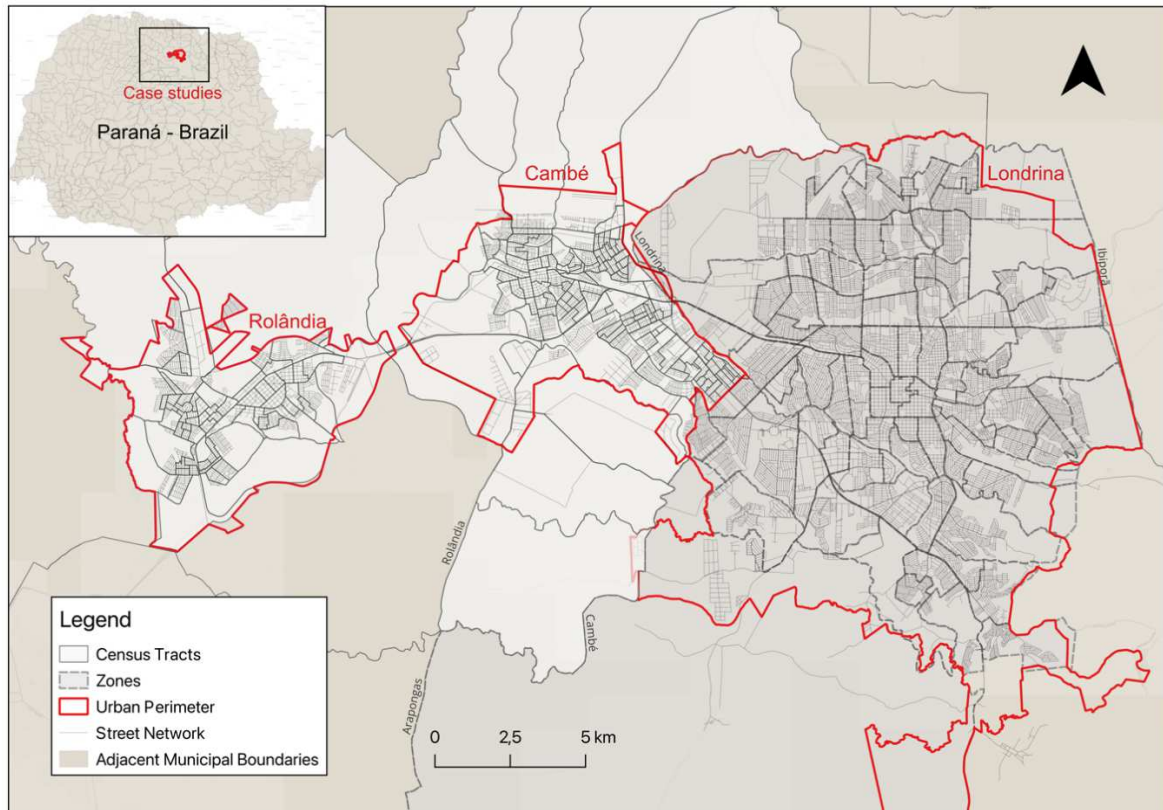
Cambé is a municipality located in the Metropolitan Region of Londrina, in the north of the State of Paraná, with a population of 105,704 inhabitants and an area of 494,692 km<sup>2</sup>,

according to the 2010 census of the Brazilian Institute of Geography and Statistics. The city of Cambé-PR was also chosen for its representativeness and data availability. Waling data was extracted from the OD survey prepared to produce the municipality's Mobility Plan in the year 2016. From questionnaires applied in the years 2015-2016, in 3.92% of the total of 32,872 permanent private households in the city, a total of 3348 were systematized where 1080 trips made by foot were successfully geocoded by the authors. The georeferencing of spatial data was also conducted by the researchers.

As a case, the city of Londrina-PR was selected, mainly due to the availability of extremely current and precise data (2020) on urban mobility. The City Hall of Londrina, through its Institute for Research and Urban Planning (IPPUL), developed in 2020, with a specialized consultancy, an OD survey, which have been made publicly available such data (<https://ippul.londrina.pr.gov.br/>). The survey was conducted in a sample of 5.131 valid households distributed across the territory in such a way as to represent the households in each zone of the city. A total of n=18635 of which were successfully systematized for this research, 4162 of these being walking trips. This is the general response variable for the statistical analyses to be conducted.

Given that the aim is not to delve into the complexities or difficulties associated with defining exposure boundaries, the term 'neighborhood' is used as a general representation of spatial exposure, despite the acknowledged associated limitations (Tobin et al., 2022). Therefore, in Rolândia, the aggregation units adopted as neighborhoods correspond to the urban census tract boundaries provided by the Brazilian Institute of Geography and Statistics (IBGE). In Londrina, the territory within the municipal boundaries was initially divided into 91 zones. These zones were produced by the city administration and the company that conducted the OD survey, and they are intended to represent aggregations of census tracts that align more closely with the existing neighborhood divisions accepted by the city administration. However, only the 83 zones within the municipal urban perimeter were considered for this research. For Rolândia and Cambé, urban census tracts provided by IBGE were adopted as units of analysis, with 72 tracts considered for Rolândia and 97 tracts for Cambé (Figure 7).

Figure 7 - Aggregation units for the three case studies – census tracts and zones.



Source: Organized by the author, 2023.

In the case of Londrina, sociodemographic information about participants is available. This allows us to assess the impact of variables such as gender, age, income, and employment status on walking behaviors, particularly in conjunction with the built environment. These factors were taken into account at this stage of the research. For Rolândia, the OD survey provides sociodemographic data for participants who reported their trips. To describe the population participating in the survey, we used the secondary data available from the Rolândia administration, focusing on factors such as age, reported gender, household composition, and family income. However, Cambé does not have available sociodemographic data for the population responding to the OD survey. Given the pressing need to moderate the influence of BE on mobility patterns using social composition data, we adopted the nominal income data from IBGE (Brazilian Institute of Geography and Statistics) as an aggregate measure. This was used to represent income by census tract of residence for those responding to the OD survey. Considering countless studies in the area of walkability and active travel, this selection of social and individual characteristics is justified, being these the most prominent and influential variables in walking levels (Cauwenberg et al., 2011; Dyck et al., 2015; Luan et al., 2019; Van

Cauwenberg et al., 2012). Table 3 summarizes sample sizes obtained from OD surveys for all case studies and the sociodemographic characteristics adopted.

Table 3 – Sample sizes for OD survey data, adopted sociodemographic variables, their classes, and data sources.

| Case study | Complete sample vs walking sample size | Adopted sociodemographic variables | Classes for sociodemographic variables                                   | Sources   |
|------------|--|------------------------------------|--|---|
| Londrina   | 18635 - 4371                           | Age                                | Child ( $\leq 12$ )<br>Adult ( $<13$ to $59<$ )<br>Elderly ( $\leq 60$ ) | Origin and Destination survey provided by city administration |
|            |  | Gender <sup>2</sup>                | Female<br>Male   |   |
|            |  | Household income                   | $\leq 2$ MW,<br>$<2$ to $5$ MW<br>$\leq 5$ MW<br>Absent                  |   |
|            |  | Employment status                  | Formal<br>Informal<br>Unemployed/Refusal                                 |   |
| Cambé      | <b>3348 - 1080</b>                     | Household income                   | $\leq 2$ MW,<br>$<2$ to $5$ MW   | IBGE, 2010  |
| Rolândia   | 2152 - 394                             | Age                                | Child ( $\leq 12$ )<br>Adult ( $<13$ to $59<$ )<br>Elderly ( $\leq 60$ ) | Origin and Destination survey provided by city administration |
|            |  | Gender                             | Female,<br>Male<br>Absent  |   |
|            |  | Household income                   | $\leq 2$ MW,<br>$<2$ to $5$ MW<br>$\leq 5$ MW<br>Absent                  |   |

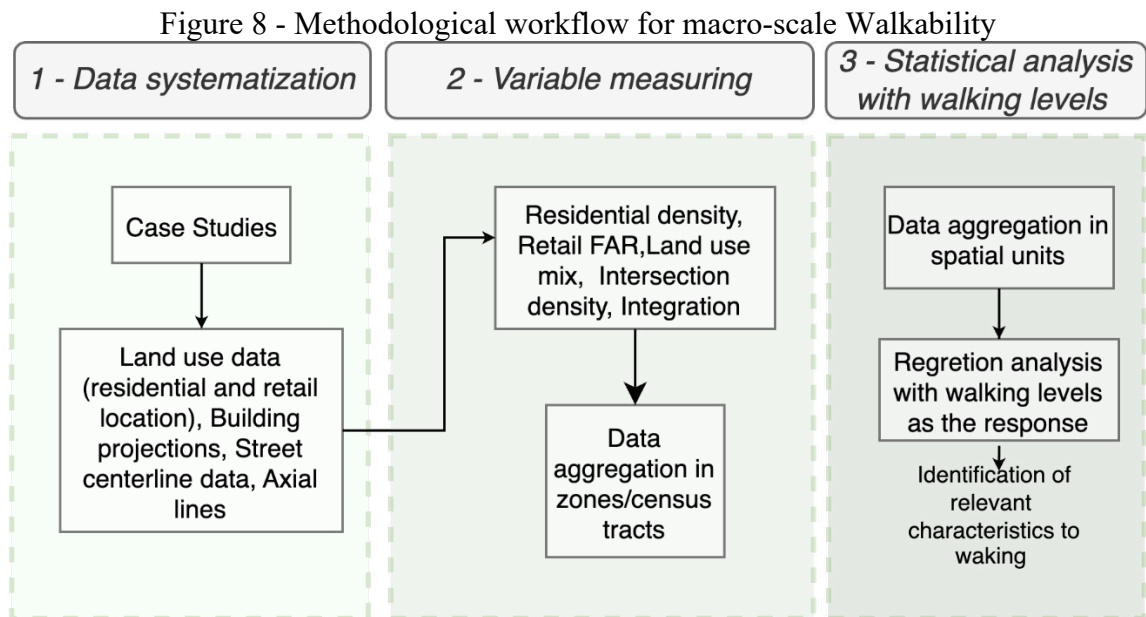
Source: The author, 2023.

### 3.2 The first methodological phase – Macro scale Walkability

The first step of the research process focuses on the macro level walkability, being further divided into 3 steps: data systematization, variable measuring, and statistical analysis with walking levels. The first stage begins with the systematization of data, where necessary information is organized for each case study. The second stage encompasses the quantification of variables identified as significant for macro-scale walkability and aggregates them into units representative of neighborhoods. The third stage involves the statistical analysis of these

<sup>2</sup> The original language used by the OD surveys refers to "sex", a term associated with biological differences between males and females. However, in this context, we focus on "gender", highlighting roles, behaviors, and activities individuals adopt, which emphasizes socially constructed differences. It's important to note that gender can address individual, cultural, institutional, and structural differences, illustrating a broader spectrum of identities and societal roles. Consequently, the data provided here primarily addresses "gender" rather than "sex" (Short SE, Yang YC, Jenkins TM, 2013).

quantifications in relation to walking behavior, identifying which characteristics of this scale are most relevant to predicting this behavior (Figure 8).



Source: Organized by the author, 2020

In the **first and second steps**, the objective variables identified as relevant in the literature review were systematized and measured according to reliable sources or gathered by researchers from the State University of Londrina (Table 4). From land use data, retail building projections, and street centerline data (that gives subsidies to the construction of axial lines for space syntax), Residential Density; Retail Floor-Area Ratio, Intersection Density, Land Use Mix (Entropy), and Space Syntax Integration were quantified per neighborhood - zones in Londrina and per census tracts in Rolândia and Cambé.

Table 4 – Macroscale walkability variables considered and data sources.

| Variables                  | Measurement methodology   | Data sources Londrina  | Data sources Cambé   | Data sources Rolândia  |
|----------------------------|---|--|--|--|
| 1 Residential Density      | Residential units per unit of area.   | Land use data provided by IPPUL  | Land use data gathered by researchers  | Land use data gathered by researchers  |
| 2 Retail Floor-Area Ratio  | Area of the retail parcels divided by the footprint of the retail building.   | Land use data provided by IPPUL and Retail building projections extracted from the Open Street Map free database | Land use data gathered by researchers and Retail building projections extracted from the Open Street Map free database | Land use data gathered by researchers and Retail building projections extracted from the Open Street Map free database |
| 3 Intersection Density     | The ratio between the number of true intersections (between three or more roads) and the extent of the zone   | Geolocated by the authors based on street centerline data provided by IPPUL                                      | Geolocated by the authors based on street centerline data provided by Cambé administration                             | Geolocated by the authors based on street centerline data provided by Rolândia administration                          |
| 4 Land Use Mix (Entropy)   | The measure of the diversity of uses present in an area unit, considering different use categories, calculated through the following formula based on Shannon (1948). | Land use data provided by IPPUL  | Land use data gathered by researchers  | Land use data gathered by researchers  |
| 5 Space Syntax Integration | Calculated through axial lines; an average integration score calculated for each zone.  | Based on street centerline data provided by IPPUL – converted into axial lines by the authors.                   | Based on street centerline data provided by Cambé administration – converted into axial lines by the authors.          | Based on street centerline data provided by Rolândia administration – converted into axial lines by the authors.       |

Source: Leão et al. (2020a), Adapted by the author, 2023.

**Residential density** was calculated as the number of residential units per zone or census tract divided by its area (B. E. Saelens, Sallis, & Frank, 2003). Land use data made available by the sources indicated in Table 4 were base for isolating such residential units. In the case of Londrina, the dataset was converted from drawing to a georeferenced database by the authors (2019). For Rolândia and Cambé, the databases were developed by the author’s research group in 2018 and 2019, respectively.

Retail parcels were isolated from the land use databases, and building footprint data was extracted as a shapefile of polygons from the OpenStreetMap dataset (available from <https://planet.osm.org/>). Then, **Retail FAR** was calculated considering the ratio between the size of retail parcel and the footprint of retail buildings per unit area.

**Intersection density** was represented by the ratio between the number of true intersections (between three or more streets) and the area of the zone or census tract (Frank et al., 2009). Intersections were mapped using street centerline in a georeferenced environment. The **Intersection density** measure was calculated considering the division of N true intersections in a unit and the area in square meters of that same unit. Zones and census tracts have their boundaries delimited by the street network; therefore, many intersections are located over boundaries of adjacent units. As a workaround for this issue, the methodology proposed by (Leão, 2019) was applied, considering the k-Nearest Neighbor classification approach where the parameter k was set to 3.

The measure of land use mix was determined using a variation of Shannon's entropy equation, which is an indicator of the level of diversity in land use distribution (Hajna et al., 2014). This measure has been commonly utilized in walkability studies, such as conducted by Frank et al. (2009), where a mix of five different land uses is considered. For the cities of Rolândia and Cambé, this particular category was chosen during the creation and data collection for the land use databases. However, for Londrina, the categories that had been previously set for the existing secondary database for land use differed. With this in mind, a modification of the category grouping was undertaken to align as closely as possible to the original classification. This adaptation is presented in Table 5, which outlines the changes made to the land use categories for Londrina, based on the original land use categories proposed by Frank et al. (2009) that were adopted for Rolândia and Cambé, and adapted from Londrina's City Hall and IPPUL's guidelines.

Table 5 – Land use category modifications for Londrina.

| Land Use Categories Proposed by Frank et al. (2009), originally adopted for Rolândia and Cambé | Adapted Land Uses Categories From Londrina's City Hall and IPPUL |
|--|--|
| 1 <b>Residential</b>   | Residential  |
| 2 <b>Retail - Service</b>  | Commerce Services/ Industrial/ Mixed                             |
| 3 <b>Institutional</b>   | Institutional Private and Public/Religious                       |
| 4 <b>Entertainment</b>   | Public open spaces   |

Source: Organized by the author, 2023.

Entropy was calculated through the following formula based on (Shannon, 1948), where  $k$  = categories of land use;  $p$  = proportion between the area of land use and the area of the census tract; and  $\ln$  = logarithm (L. D. Frank et al., 2009). The resulting values are normalized between 0 and 1, where 0 would indicate only one use in a given area, and 1 would indicate a complete and equal distribution of the five uses.

$$- \sum k = \left( \frac{pk \times \ln pk}{\ln N} \right)$$

The land use data from the Urban Planning Institute of Londrina, sourced from a 2018 survey, categorizes land use across a broad spectrum. However, this dataset does not provide information on the number of units per plot, only identifying the prevailing class for each lot. This lack of granular information can lead to substantial difficulties, especially when measuring the variable of mixed land use. To address this problem, we initially explored the open data from OpenStreetMap (OSM), which offers in-depth details regarding land use of plots and the number of storeys in buildings across various global cities. Yet, for Londrina, only around 4% of the city's total plots were accompanied by relevant usage and floor count information. Consequently, OSM proved to be a rather inefficient information source for potentially supplanting the data provided by Londrina's Urban Planning Institute.

As an alternative, we examined data from the Brazilian Institute of Geography and Statistics (IBGE), within the data from the National Registry of Addresses for statistical purposes. This dataset provides nationwide information on households, and other establishments, inclusive of details on land use, number of uses per plot, and floors in a highly specific manner. Yet, this data, dating back to 2010 and released in tandem with the 2010 Census, carries a discrepancy of at least 13 years, which might affect its relevance.

In order to evaluate the potential of this data from the National Address Registry for measuring variables of walkability - based on land use data for predicting walking propensity - it was necessary to systematize this data concerning land use and the number of floors per plot for the city of Londrina specifically. To facilitate this, from the available information from the IBGE download page was parsed using a Python script and processed it within an R script to calculate entropy following the Shannon equation. The smallest possible aggregation unit of the National Address Registry data, block faces, was used for this aggregation. For the calculation of Shannon's entropy based on the data from IBGE's National Registry of Addresses

for statistical purposes, four land use categories were taken into account: commercial, residential, institutional, and hospitality or hotels. Subsequent to the block entropy calculation for Londrina, the results were extrapolated to zones using an aggregation of this information by average. This approach allowed assigning the entropy value to the corresponding zone. The thus obtained information was then employed for a comparative analysis with recent data provided by Londrina's municipality and its Urban Planning Institute.

**Space Syntaxes' Integration** was calculated based on street centerline data provided according to sources shown on Table 4 and adapted to represent segments imported into the QGIS software, a free and open-source Geographic Information System (QGIS DEVELOPMENT TEAM, 2020). Through the Space Syntax toolkit (Gil, 2020) and depthmapX (DepthmapX DEVELOPMENT TEAM, 2017), the integration measure was calculated for each street segment within a 1200-meter radius, equivalent to a 15-minute walk (Kronenberge and Saboya, 2019). The urban space's configurational characterization was conducted through an angular analysis of segments. Subsequently, the average integration of the streets belonging to each zone was calculated to assign an average integration per zone.

In Space Syntax, urban space is divided into spatial units known as axial lines. These are the largest straight lines capable of covering a whole system of public spaces, possibly being considered as segments – lines with the highest possible straight visibility (Hillier and Hanson, 1984). The relationships between such elements of a system can be analyzed using the Integration measure, which indicates how “deep”, that is, integrated, a network element is from all other elements (Hillier et al., 1993). The average depth of the  $i$ -th element can be calculated by the summation of the depths of the  $k$  elements in relation to it, divided by the total number of elements minus one:

$$MD_i = \frac{\sum_{j=1}^k d_{ij}}{(k-1)}$$

MD<sub>i</sub> = Average Depth;  
d<sub>ij</sub> = Depth of line  $j$  in relation to line  $i$ ;  
 $k$  = total number of system components

The most connected lines, closest to all others in a system, are called integrated. On the other hand, the most distant ones are called segregated.

**For the third and final step** from the methodological workflow for Macro-scale Walkability, regression analyses were conducted. Walking levels were a response variable for walking or not (as a binary response) considering the sample of trips for each of systematized OD surveys. The built environment characteristics presented in Table 4 were taken as

explanatory variables attributed to each trip by considering the respondent's originating zone. The sociodemographic characteristics from Table 3 were considered as covariates. Multilevel regression modeling was conducted to examine walkability variables' associations with the likelihood of walking while controlling the clustering pattern of participants' walking behaviors from the same zone. The Interclass Correlation Coefficient (ICC) is utilized to demonstrate the extent to which the observed variance in walking behavior is linked to variations at the census tract or zone level. This underscores the significance of data aggregation at this specific spatial unit. The intraclass correlation coefficient, on the other hand, measures the correlation between individuals who are grouped within the same context. Essentially, it quantifies the proportion of the total unexplained variation in an outcome that arises due to differences between these contexts (Leyland AH & Groenewegen PP., 2020). Continuous variables were standardized as z-scores and average z-scores.

The feature that distinguishes a multilevel regression model from a classical one is modeling the variation between groups, in this case – zones or census tracts. For all regression analyses proposed here, from previous phases and this final modeling, in estimating walking with built environment factors, there are two levels: level-1 units (walking trips) are clustered within level-2 units (zones/census tracts).

Individuals that reside in the same place tend to be similar. This can lead to inconsistent results by violating the ordinary regression models' assumption that individual observations (walking trips in this case) should not affect or be affected by other observations (Gelman & Hill, 2007; Kim et al., 2014). Multilevel modeling deals with this spatial autocorrelation issue by accounting for both the variability among zones (level-2 units) and the variability among trips (level-1 units).

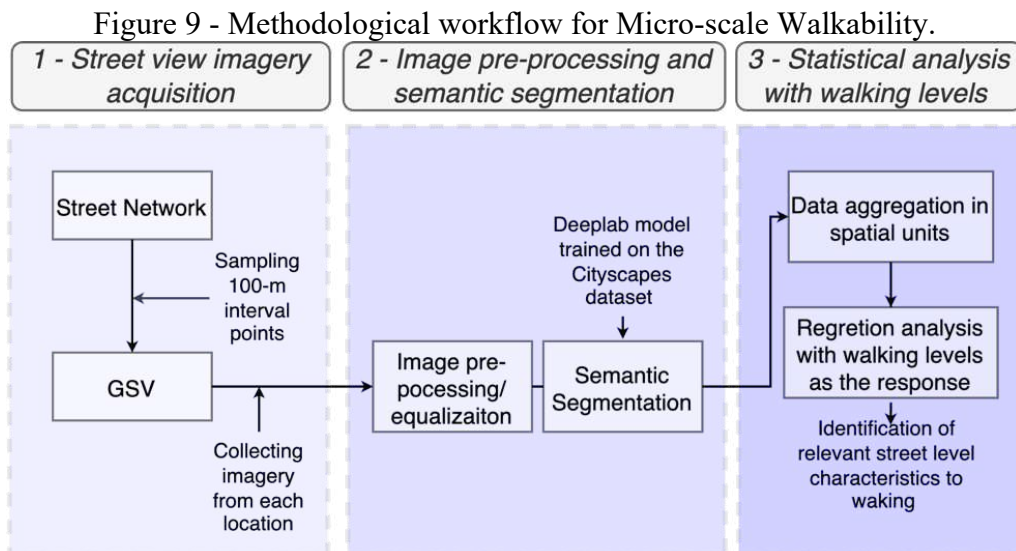
Such statistical modeling was conducted in R (R Core Team, 2020), specifically with the 'lme4' package and its 'glmer' function, where both fixed effects (independent built environment variables) and random effects (zones) are specified in the model formula. Odds Ratio (OR) and 95% confidence intervals (CI) were reported for interpretation. The correlation among predictors was verified through the Variance Inflation Factor (VIF) with the 'car' package and the 'vif' function. VIFs smaller than 2 were considered as the indication that multicollinearity was not present. When a model includes multiple factors correlated with each other rather than just the response variable, multicollinearity may exist. The general criterion is that VIFs exceeding four warrant further investigation into multicollinearity, while VIFs exceeding 10 require correction (O'Brien, 2007).

Three different models for each of the case studies were constructed for discussion analysis purposes: one considering only BE variables as explanatory variables; a second (called a complete model) with both be variables and sociodemographic; and the best-fitted model including only the significant variables for the model. For model comparison, the Akaike Information Criterion (AIC) was used as a goodness of fit metric.

### 3.3 Second Phase Microscale Walkability

The second phase of the research process focuses on micro-level walkability, further divided into three steps: street view image acquisition, image pre-processing and semantic segmentation, and statistical analysis with walking level (Figure 9).

**Initially, in the first step of the acquisition of images**, sample of points were distributed every 100 meters through the street network using the ArcGis 10.6 software. The Data Management toolbox used was the *Generate Points Along Lines* tool, allocating points along lines at fixed intervals over street centerline data provided city administrators of the three case studies. The selected metric interval was adopted according to literature tendencies, that reflect many possible intervals for sampling without consensus - for example, 20 m (Seiferling et al., 2017), 50 m (LU, 2018), and 100 m (Helbich et al., 2019) also due to processing capacity limitations.

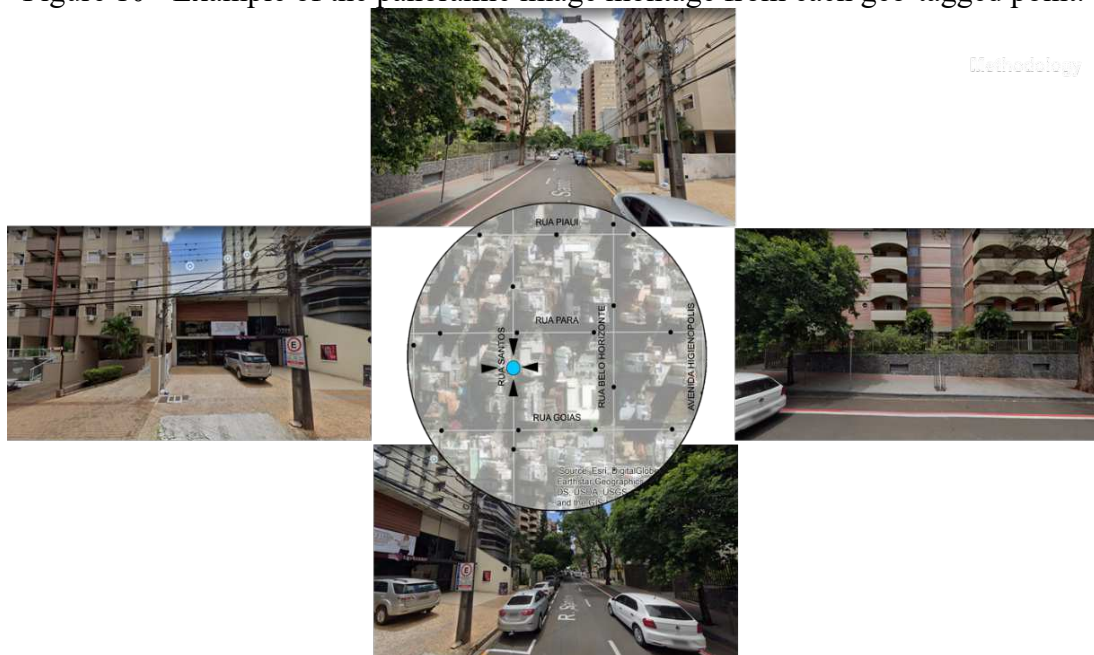


Source: Organized by the author, 2023.

The images were acquired using the Google Street View API (Application Programming Interface), representing the pedestrian level's cityscape. Through the interface, images were obtained from the geo-tagged sample points. Such a process begins with a metadata request

with the information from the images. Firstly, the locational parameters were indicated through latitude and longitude from the previously sampled points. Secondly, the images' source was defined as outdoor, as the procedure is aimed at streetscape images. From this metadata, the request to obtain the images needs an image size parameter; a field of vision parameter that was standardized here at 0, 90, 180, and 270 degrees to form a panorama image (Figure 10); the horizontal heading that indicates the compass direction in which the camera is pointing; and the vertical camera direction, that specifies the vertical angle of the camera. The construction of the code for obtaining the images was implemented with the Python programming language.

Figure 10 - Example of the panoramic image montage from each geo-tagged point.



Source: Organized by the author, 2023.

After collecting samples, **the second step was the pre-processing of the images**, aiming to improve the quality of the predictions produced by the model and avoid overfitting due to uninformative characteristics. Considering that the GSV images were acquired at different times of the day and in different climatic conditions, image processing is necessary to prevent lighting and saturation highlighted by the model and recognized as information related to the neighborhood (Leão et al., 2020). Further, general filtering of the dataset was conducted to remove erroneous images with blurs or obstructions.

To discern the compositional elements in the images from each sampled point, the SegFormer model was employed. Given the rapid advancements in semantic segmentation models and the crucial role that high-quality predictions play in data analysis, this model, which is the state-of-the-art at the time of writing, by Xie et al. (2021) was adopted. The pretrained

model available used to infer the classes of the pixels from collected images. SegFormer is a lightweight Transformer-based semantic segmentation model that achieves 84.0% mean pixel intersection-over-union (mIOU) on Cityscapes dataset. Moreover, it is a lightweight Transformer-based semantic segmentation model that achieves relevant results on Cityscapes-C, a corrupted version of Cityscapes dataset, stating its generalization and robustness. Thus, according to the literature benchmarks, SegFormer demonstrated to be a more reliable model than its alternatives, such as DeeplabV3+ and PSPNet. The publicly available version of SegFormer fine-tuned on the Cityscapes dataset was adopted<sup>3</sup>.

This methodological choice stems from the model's recency and superiority over the DeepLab v3+ for example, adopted in the initial stages of this research. Further, it has been shown to have high performance when trained in the chosen training dataset. The Cityscapes Dataset (Cordts et al., 2016) was adopted according to literature tendencies. It is an image dataset with annotations of streetscape segments defined in seven larger groups (i.e., flat, human, vehicle, construction, object, nature, sky, and void). The dataset provides 19 classes for training (Table 6)

These classes are not necessarily equivalent to any pre-existing micro-scale audits; however, they represent significant elements that might be present in many of these tools. The procedure proposed here is not thought out to be equivalent to any Especific tool. Alternatively, this is a proposition of an alternative methodology that automatically identifies BE components. Past this consideration, it is possible to use the categorization proposed by Moudon and Lee (2003) in the classes available for the Cityscapes Dataset (Table 6). Spatiophysical elements are here represented by the Building and Sidewalk classes as general infrastructure. Spatiopsychosocial elements that influence people's perception of the environment are represented, for example, Vegetation and Sky (enclosure). Spatiobehavioral elements that directly influence people's behavior in urban environments are represented by identifying the presence of Cars and Traffic Signs.

---

<sup>3</sup> <https://huggingface.co/nvidia/segformer-b5-finetuned-cityscapes-1024-1024>

Table 6 - List of annotated classes in the Cityscapes dataset.

| <i>Group</i>        | <i>Class</i>  | <i>Description</i>   | <b>**</b>         |
|---------------------|---------------|--|-------------------|
| <i>Flat</i>         | Road          | An area where cars usually drive, e.g., lanes, directions, and streets. Areas only delimited by markings from the main road are also roads, e.g., bicycle lanes, roundabout lanes, or parking spaces. This label does not include curbs.   | Spatio-physical   |
|                     | Sidewalk      | An area located at the side of a road and delimited from the road by some obstacle for pedestrians or cyclists, e.g., curbs or poles (perhaps small). This label includes a possibly delimiting curb, traffic islands (the walkable part), or pedestrian zones (where usually cars are not allowed during daytime).                              |                   |
| <i>Construction</i> | Building      | Building, skyscraper, house, bus stop building, garage, carport.   |                   |
|                     | Wall          | Individual standing wall. Not part of a building   |                   |
|                     | Fence         | Fence including any holes.   |                   |
| <i>Nature</i>       | Vegetation    | Tree, hedge, all kinds of vertical vegetation. Plants attached to buildings are usually not annotated separately and are labeled “building” as well. If growing at the side of a wall or building, it is marked as vegetation if it covers a substantial part of the surface (more than 20%)   |                   |
|                     | Terrain       | Grass, all kinds of horizontal vegetation, soil, or sand. This label includes a possibly delimiting curb.  |                   |
| <i>Sky</i>          | Sky           | Open sky, without tree leaves. Includes thin electrical wires visible in skyscape.   |                   |
| <i>Human</i>        | Person        | A human walking, standing, or sitting on the ground, on a bench, or a chair. This class includes toddlers and someone pushing a bicycle or standing next to it with both legs on the same side of the bicycle. This class also includes anything carried by the person, e. g., backpack, but not items touching the ground, e.g., trolleys.      |                   |
| <i>Vehicle</i>      | Rider         | A human using some device to move, e.g., riders/ drivers of bicycles, motorbikes, scooters, skateboards, horses, roller-blades, wheel-chairs, road cleaning cars, cars without a roof. Note that a visible driver of a car with a roof can be seen only through the window. Since holes are not labeled, the human is included in the car label. | Spatio-behavioral |
|                     | Car           | Automobile, jeep, SUV, a van with continuous body shape, caravan, but no trailers.   |                   |
|                     | Truck         | Truck, box truck, pickup truck, including their trailers. the back part/loading area is physically separated from the driving compartment.   |                   |
|                     | Bus           | Bus for 9+ persons, public transport, or long-distance transport   |                   |
|                     | Train         | Vehicle on rails, e.g., tram, train.   |                   |
|                     | Motorcycle    | Motorbike, moped, scooter without a driver (a rider—see above)   |                   |
| <i>Object</i>       | Bicycle       | A bicycle without a driver (a rider—see above).  |                   |
|                     | Pole          | Small, mainly vertically oriented pole or having a diameter (in pixels) of at most twice the diameter of the pole, e.g., sign pole, traffic light poles, streetlights  |                   |
|                     | Traffic sign  | A sign without a pole, showing information of the driver/cyclist/pedestrian in an everyday traffic scene, e.g., traffic signs, parking signs, direction signs. This label counts only the front side of a sign containing  |                   |
|                     | Traffic Light | A traffic lightbox without its poles.  |                   |

\*\* Aspects of Behavior-Environment

Source: Organized by the author, 2020. Based on (Cordts et al., 2016), Nagata et al. (2020) and Moudon; Lee (2003)

Every pixel of all GSV images was classified into the established segments and quantified. This procedure's output was a .csv table indicating the absolute count of pixels, a ratio, and a z-score of each segment class for each image. The semantic segmentation results were aggregated in zones by averaging all sampled points' corresponding semantic category scores.

In the third and final phase of this methodological process for micro-scale Walkability, statistical analyses were conducted in conjunction with the levels of walking activity. The initial analysis aimed to explore the data using the average aggregated values from each category and the total number of walking trips per zone or census tract. At this point in the methodology, considering the importance and the widely recognized scientific differences between the motivators of recreational walking and utilitarian walking (i.e., for transportation or leisure) (B. Kang et al., 2017), an initial approach was undertaken to differentiate between these two types of activities. The activities were divided into total walking, which represented the absolute number of walks conducted per zone or census tract, and total recreational and utilitarian walking, measured by the absolute number per aggregation unit. This categorization aimed to emphasize the distinction between walks taken for different purposes and ensure a comprehensive understanding of walking behavior. Spearman's rho correlation coefficients were obtained with the due Hypothesis testing for the coefficients' significance at the  $\alpha=0,05$  level. Such analysis was conducted in R (R Core Team, 2020) with the 'cor' and 'cor.mtest' functions.

**A second analysis aiming to verify the micro-scale-built environment's predictive power** over walking was conducted, controlling for sociodemographic characteristics. *By doing so, a view of the problem from the individuals' perspective about the urban spaces can be seen.* Considering the high number of semantic segmentation categories for statistical modeling, feature selection was applied. The walking data was not aggregated; therefore, the sample is related to the total number of trips and their respective zone's micro-scale characteristics. The more contemporary wrapper method based on the random forest model Boruta was selected (Kursa & Rudnicki, 2010).

All such analyses were conducted in R (R Core Team, 2020). Feature selection was constructed with the 'Boruta' package. The 'lme4' package and its 'glmer' function were used for the multilevel regression modeling. OR and 95% CI were reported for interpretation. The correlation among predictors was verified through the Variance Inflation Factor (VIF) with the 'car' package and the 'vif' function.

Three different models for all case studies were constructed for discussion analysis purposes: one considering only BE variables as explanatory variables; a second (called a complete model) with both BE variables and sociodemographic information on gender, income and age; and the best-fitted model including only the significant variables for the model. For model comparison, the AIC was used as a quality of fit metric

Spatial distribution mapping was conducted through *Hotspot analyses* (Ord & Getis, 1995).

The distribution of the most positively relevant semantic segmentation categories in the regression model could be analyzed regarding where each of the categories is significantly spatially grouped (i.e., the spatial concentration of these characteristics is unusual in the system/urban space).

### **3.4 Third Phase Comprehensive Walkability modeling**

Finally, a regression model was developed for each of the case studies. During the earlier research stage, which included the preliminary board examination, full modeling wasn't carried out as both theoretical and operational aspects still required fine-tuning. The objective of this third phase was to examine the relationship between micro and macro level characteristics and walking activity, using multilevel regression modeling (Austin et al., 2001). Analyses were conducted in R (R Core Team, 2020). The 'lme4' package and its 'glmer' function were used for the multilevel regression modeling. OR (OR) and 95% confidence intervals (CI) were reported for interpretation. The correlation among predictors was verified through the Variance Inflation Factor (VIF) with the 'car' package and the 'vif' function.

Relevant multiscale variables and sociodemographic factors were comparatively analyzed for models in three different sized cities. All three methodological steps were implemented across all case studies, leading to the development of representative models for each city. These models can subsequently be used to verify the research thesis. It's worth reiterating that these models were utilized in this research to describe and comprehend the problem or phenomenon at hand. However, at this stage, they cannot be viewed as a commercial or practical representation of an operational urban policy product.

#### 4 RESULTS AND DISCUSSIONS

The following results are presented **first introducing the characteristics of the secondary data based on origin-destination surveys and socio-demographic characteristics** of the samples used for each case study are presented. This is followed by the presentation of the sample characteristics in relation to the mode of transportation in the origin-destination survey, as well as the graphical representation of the aggregation of the number of walking trips in the aggregation units used - census sectors or zones. An analysis of the average distance of walking trips based on a histogram is also presented.

Subsequently, **a description of the data used and the results of the multilevel regression model considering only macro-scale variables** and individual variables for predicting walking are presented for each city: Rolândia, Cambé, and Londrina.

Following this, **information related to the systematization and description of results of walkability analyses at the micro-scale are presented**. This starts with the presentation of results for point sampling for the extraction of images from Google Street View. An initial statistical analysis based on a non-parametric correlation between walking trips for leisure and transportation is then presented, providing a deeper understanding of the differences between the reasons for walking trips and the micro-scale variables. Results of statistical modeling of the multilevel regression are presented next, primarily the multilevel regression model considering only micro-scale variables and individual variables for predicting walking for each city.

Finally, the last chapter of **descriptively presents results and discusses the findings of res multilevel modeling**, i.e., at the micro and macro levels for the three cities where multilevel logistic regression models were developed for each city. These models consider individual variables and relevant variables from the micro and macro scales together to predict the choices of walking in each of the contexts - small, intermediate, and medium-sized cities, with approximately 70,000, 100,000, and 500,000 inhabitants respectively.

#### 4.1 Results for Rolândia

Analysis conducted in Rolândia -PR considered the n=2152 valid trips conducted by any modal extracted from the municipal OD survey were considered the study sample. A total of n=394 were walking trip. This is the general response variable for the statistical analyses conducted. The descriptive statistics of the sociodemographic characteristics of the individuals who reported the trips are shown in Table 7. When it comes to gender, the sample is reasonably balanced. However, for Age, the ‘adult’ category is heavily oversampled, which might affect results. Income data is also very unbalanced, underrepresenting higher income classes.

Table 7 - Sociodemographic characteristics of the sample for Rolândia (n = 2152).

| Sociodemographic Variables | Count | Percentage (%) |
|----------------------------|-------|----------------|
| <b>Age</b>                 |       |                |
| Child ( $\leq 12$ )        | 244   | 11.34          |
| Adult ( $< 13$ to $59 <$ ) | 1630  | 75.74          |
| Elderly ( $\leq 60$ )      | 278   | 12.92          |
| <b>Gender</b>              |       |                |
| Female                     | 1134  | 52.69          |
| Male                       | 979   | 45.49          |
| Absent                     | 39    | 1.81           |
| <b>Household income</b>    |       |                |
| $\leq 2$ MW                | 1095  | 50.88          |
| $< 2$ to $5$ MW $<$        | 140   | 6.50           |
| $\leq 5$ MW                | 12    | 0.56           |
| Absent                     | 905   | 42.05          |

Source: The author, 2023.

In the analysis of the sample distribution and its sociodemographic characteristics, it can be identified firstly that, among individuals who identified as female, a higher percentage of trips were made on foot (22.57%). Secondly, there was a higher proportion of walking trips among individuals from lower income classes, indicating a potential correlation between income level and mode of transportation. Furthermore, it was observed that individuals who refused to answer or did not provide a response also had a higher percentage of walking trips. Lastly, both children and the elderly exhibited higher percentages of walking trips, suggesting age as a factor influencing the choice of transportation mode (Table 8)

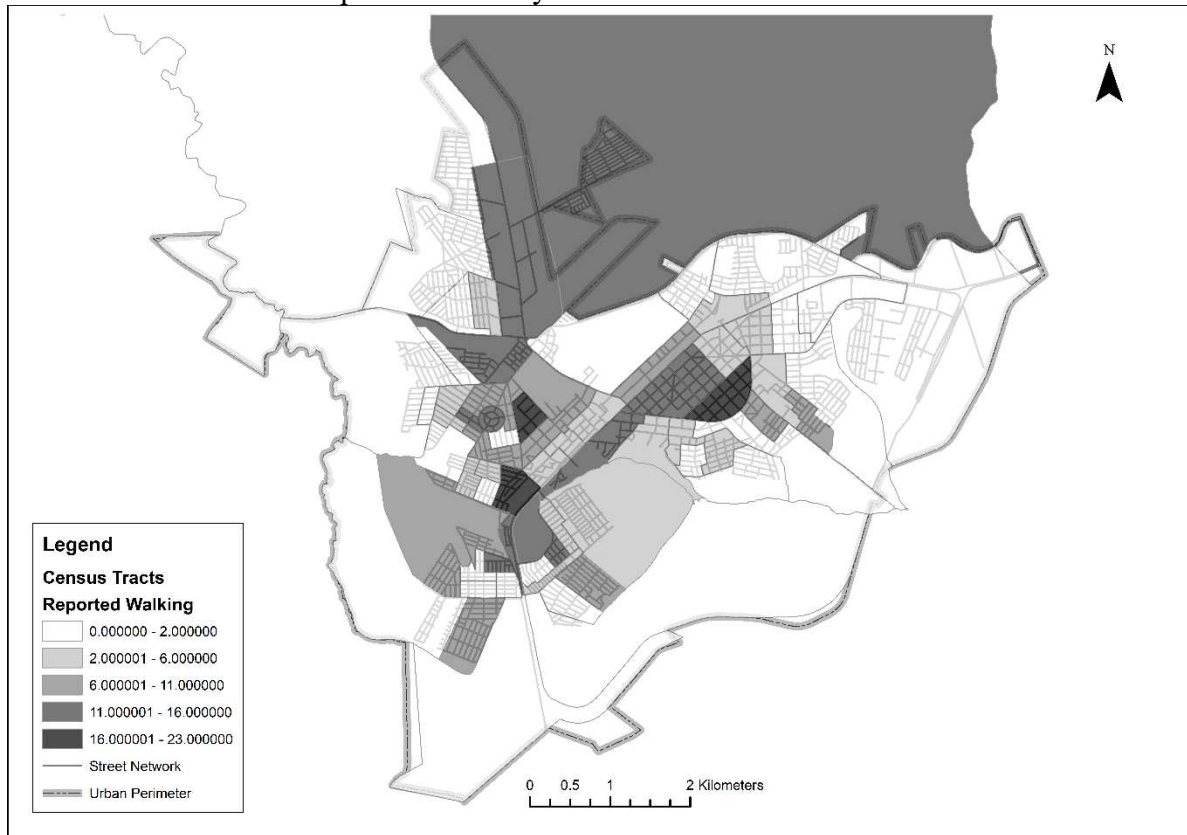
Table 8- Sociodemographic characteristics of the sample versus walking or not in Rolândia (n = 2152).

| Sociodemographic Variables          | Modal |       |      |       |
|-------------------------------------|-------|-------|------|-------|
|                                     | Other | %     | Walk | %     |
| Child ( $\leq 12$ ) (n=244)         | 171   | 70.08 | 73   | 29.92 |
| Adult ( $< 13$ to $59 <$ ) (n=1630) | 1396  | 85.64 | 234  | 14.35 |
| Elderly ( $\leq 60$ ) (n=278)       | 196   | 70.50 | 82   | 29.50 |
| Female (n=1134)                     | 878   | 77.42 | 256  | 22.57 |
| Male (n=979)                        | 848   | 86.62 | 131  | 13.38 |
| Absent (n=39)                       | 37    | 94.87 | 2    | 5.13  |
| $\leq 2$ MW (n=1095)                | 938   | 85.66 | 157  | 14.34 |
| $< 2$ to $5$ MW $<$ (n=140)         | 135   | 96.43 | 5    | 3.57  |
| $\leq 5$ MW (n=12)                  | 12    | 100   | 0    | 0     |
| Absent (n=905)                      | 678   | 74.92 | 227  | 25.08 |

Source: The author, 2023.

Figure 11 illustrates the distribution of the absolute number of trips taken per census tract in the city of Rolândia. It reveals a relatively uniform pattern without significant peripheralization of high travel levels. Particularly in the Eastern and extreme Northwest regions of the city there are fewer trips, mainly due to a potential lack of sufficient population density. This observation can possibly be attributed to their recent development, however it is worth noting that in smaller cities, pedestrian trips tend to concentrate more in central areas and denser residential neighborhoods, as depicted in the figure. The darker areas representing the higher quintile of values are in central regions characterized by a higher mix of residential and commercial uses.

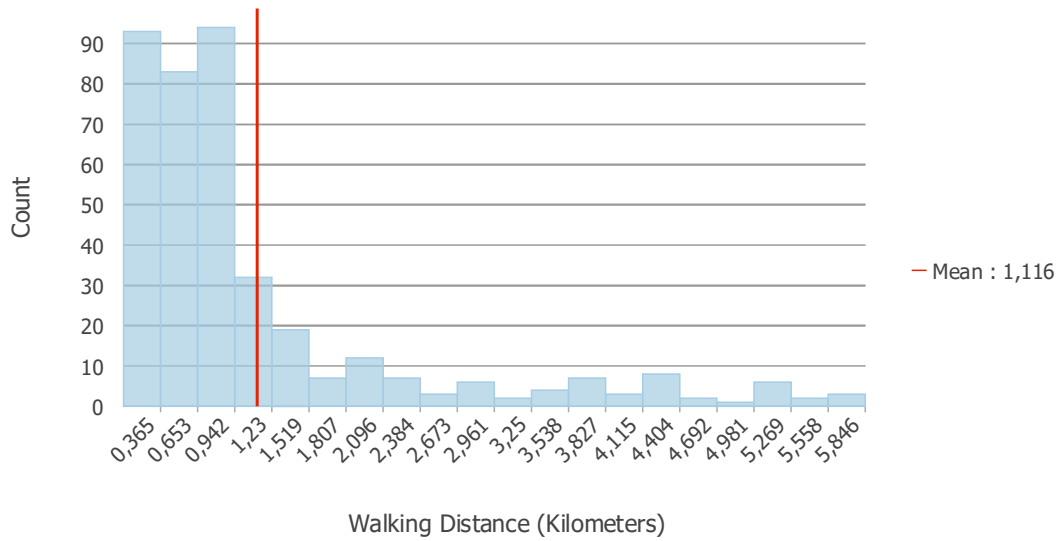
Figure 11 - Representation of the aggregate absolute number of walking trips conducted by census tracts in Rolândia.



Source: The author, 2023.

The histogram (Figure 12) illustrates the distribution of walking distances for the  $n=394$  walking trips conducted in Rolândia. The mean walking distance for these trips is 1.116 kilometers. Most of the trips fall within the shorter distances, with a significant concentration of trips ranging from 0.365 to 0.942 kilometers. As the distance increases, the frequency gradually declines, reaching its lowest point for distances near the upper end of the range (5.84 kilometers).

Figure 12 – Histogram of walking distances of the n=394 walking trips conducted in Rolândia.



Source: The author, 2023. Based on data provided by Rolândia’s administration (2018).

#### 4.1.1 Macroscale Walkability for Rolândia

Data on the location of residential lots were utilized to calculate residential density by census tracts. Information pertaining to the location of commercial lots and the projected size of the buildings on those lots were employed to determine the ratio between the built-up area and the designated commercial area. The locations of actual road intersections were identified by analyzing the city's street network and vertices of crossing in between lines/streets, which facilitated the calculation of intersection density. To quantify entropy using Shannon calculations, the researchers from the *Design Ambiental Urbano* Research Group categorized land uses into five distinct categories according to proposed by Frank et al. (2009). Lastly, for evaluating the measure of integration, the road network, configured as axial lines, served as the fundamental database (Figure 13).

Figure 13 - Macro variable for Rolândia. Respectively: residential parcels, built area (building footprint) for commercial purposes, intersections, land use data, street integration calculations from space syntax.



Source: The author, 2023.

The ICC of all modelling for the city of Rolândia is of 0.47 for the null model, indicating that approximately 47% of the variability in the odds of walking versus not walking can be attributed to differences between census tracts. The significant proportion of variance accounted for by census tracts emphasizes the relevance of considering the contextual characteristics and built environment factors within specific areas when analyzing walking behavior in Rolândia.

Three models were developed to construct the logistic regression model for the city of Rolândia (Table 9). The initial model included solely environmental variables, namely mixed land use, residential density, intersection density, the ratio of built-up area to commercial area, and the integration value within a radius of 1,200. Subsequently, a comprehensive model was constructed, encompassing all these environmental variables, along with socio-demographic variables representing individual factors, such as age, gender, and household income. The maximum VIF is approximately 4, a relatively high value, but still within the limits of safety, therefore the values obtained can be considered valid despite having this higher indication.

The categories of individual variables are determined based on the possibilities within the limitations of the utilized secondary data. Age is divided into three categories: the reference category comprises children aged equal to or younger than 12, followed by adults aged between 13 and 59, and finally, the elderly aged equal to or older than 60. The gender variable is categorized into three classes: the reference class corresponds to the reported female gender, followed by a second class representing the male gender, and a third class representing missing information. In relation to the household income variable, four classes are identified: the reference class represents lower income, defined as equal to or less than two minimum wages (MW) of family income. This is followed by a second class representing an income between two and five MW of family income, a third class representing an income equal to or greater than five MW, and finally, a class representing missing or unavailable responses.

Finally, the optimal model was identified through the employment of model comparison using AICs. The comparison revealed that all individual variables remained significant in the final model, while only integration retained relevance in predicting whether trips were undertaken by walking or utilizing an alternative transportation mode. Although the importance of the OR pertaining to an individual residing in an area with higher integration versus lower integration is acknowledged within the model, its interpretability is limited as the OR equals one. Therefore, this model, which aims to predict walking behavior binarily, utilizing macro-scale information in conjunction with individual variables such as age, gender, and household income, can be interpreted based solely on its individual variables.

Regarding age, it is identified that, with children as the reference category, adults have a lower OR of walking at 0.532, while the elderly have a significantly higher OR of walking compared to using other modes, with an OR of 1.68. Concerning gender, with the female class as the reference, individuals reporting male gender have approximately half the odds of walking

compared to using other modes. Regarding income, it is observed that individuals with higher income have progressively lower odds of engaging in walking behavior when compared to the lowest income class of two or fewer MW. Individuals with an income between two and five MW have a 36% lower chance of engaging in walking.

Table 9 - Multilevel logistic regression analysis for the relationship between the choice of walking, macro/macro-built environment, and individual factors (n = 2152).

| Model Predictor   | (1) BE variables                     |                |            | (2) Complete model                    |                 |                | (3) Best fitted model                    |                |                |
|---|--------------------------------------|----------------|------------|---------------------------------------|-----------------|----------------|--|----------------|----------------|
|   | OR                                   | 95% CI         | p-value    | OR                                    | 95% CI          | p-value        | OR                                       | 95% CI         | p-value        |
| <b>Individual</b>   |                                      |                |            |                                       |                 |                |  |                |                |
| <b>Age</b>  |                                      |                |            |                                       |                 |                |  |                |                |
| Child ( $\leq 12$ ) -   | -                                    | -              | -          | 1.00*                                 | -               | -              | 1.00*                                    | -              | -              |
| Adult ( $< 13$ to $59 <$ )                                      | -                                    | -              | -          | 0.52                                  | 0.355<br>0.760  | $< 0.001^{**}$ | 0.532                                    | 0.365<br>0.773 | $< 0.001^{**}$ |
| Elderly ( $\leq 60$ )   | -                                    | -              | -          | 1.704                                 | 1.036<br>2.801  | $< 0.001^{**}$ | 1.68                                     | 1.031<br>2.745 | $< 0.001^{**}$ |
| <b>Gender</b>   |                                      |                |            |                                       |                 |                |  |                |                |
| Female  | -                                    | -              | -          | 1.00*                                 | -               | -              | 1.00*                                    | -              | -              |
| Male  | -                                    | -              | -          | 0.493                                 | 0.379<br>0.640  | $< 0.001^{**}$ | 0.502                                    | 0.387<br>0.649 | $< 0.001^{**}$ |
| Absent  | -                                    | -              | -          | 0.074                                 | 0.013<br>0.332  | $< 0.001^{**}$ | 0.080                                    | 0.018<br>0.354 | $< 0.001^{**}$ |
| <b>Household income</b>   |                                      |                |            |                                       |                 |                |  |                |                |
| $\leq 2$ MW   | -                                    | -              | -          | 1.00*                                 | -               | -              | 1.00*                                    | -              | -              |
| $< 2$ to $5$ MW $<$   | -                                    | -              | -          | 0.355                                 | 0.137<br>0.921  | $< 0.05^*$     | 0.36                                     | 0.140<br>0.918 | $< 0.05^*$     |
| $\leq 5$ MW   | -                                    | -              | -          | 0.000                                 | 0.000 Inf       | 0.932          | 0.000                                    | 0.000          | 0.975          |
| Absent  | -                                    | -              | -          | 2.100                                 | 1.576<br>2.804  | $< 0.001^{**}$ | 2.060                                    | 1.549<br>2.745 | $< 0.001^{**}$ |
| <b>Built environment (cont. var/ average z-scores per zone)</b> |                                      |                |            |                                       |                 |                |  |                |                |
| Land use Mix (entropy)  | 0.826                                | 0.572<br>1.193 | 0.3079     | 0.823                                 | 0.584<br>1.158  | 0.262          | -  | -              | -              |
| Residential Density   | 1.171                                | 0.738<br>1.857 | 0.5023     | 1.25                                  | 0.809<br>1.915  | 0.318          | -  | -              | -              |
| Intersection Density  | 1.030                                | 0.670<br>1.583 | 0.8915     | 0.954                                 | 0.637<br>1.427  | 0.818          | -  | -              | -              |
| Retail FAR  | 0.867                                | 0.627<br>1.198 | 0.3865     | 0.842                                 | 0.6211<br>1.142 | 0.268          | -  | -              | -              |
| Integration r1200   | 1.009                                | 1.000<br>1.016 | $< 0.05^*$ | 1.005                                 | 1.001<br>1.008  | $< 0.05^*$     | 1.000                                    | 0.999<br>1.010 | $< 0.05^*$     |
| <b>Model Fit</b>  | <b>AIC = 1924.7, logLik = -955.4</b> |                |            | <b>AIC = 1804.7, logLik = -888.33</b> |                 |                | <b>AIC = 1800.1***, logLik = -890.06</b> |                |                |

Note!: \*=1,00: reference category; CI 95%: 95% Confidence Interval. \*\* significative at the  $\alpha=0,05$  level.

\*\*\* ANOVA's Pr(>Chisq) at  $\alpha=0.05$

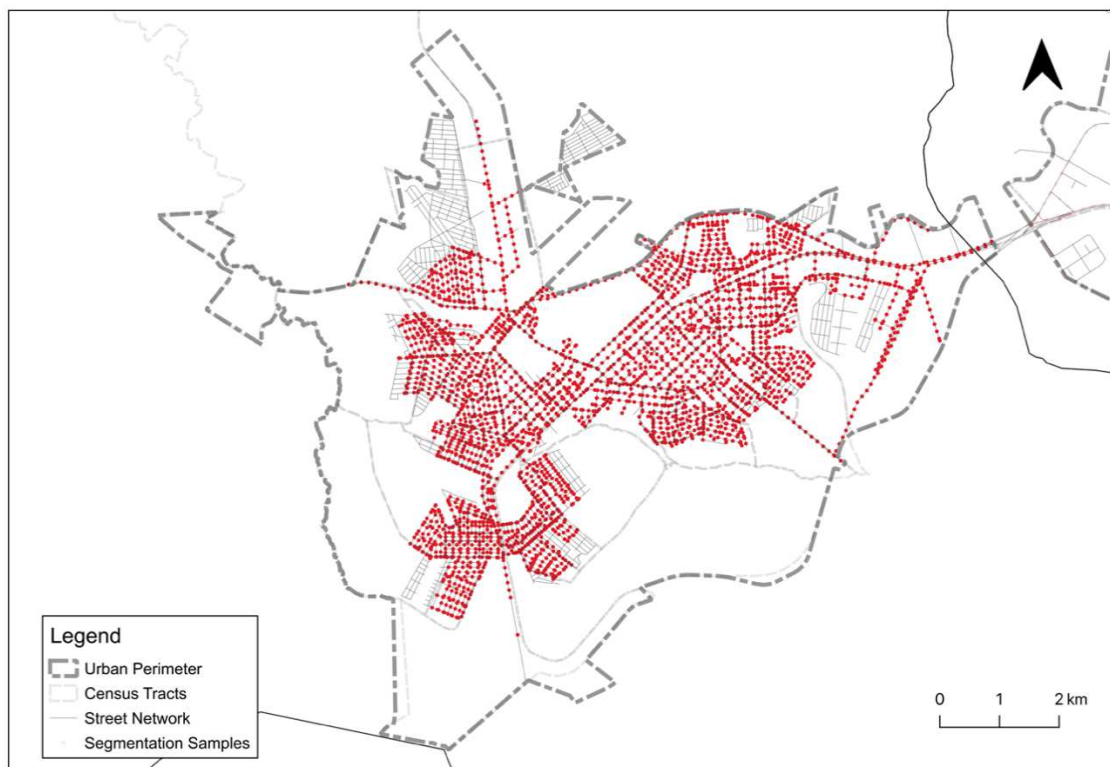
Source: The author, 2023.

#### 4.1.2 Microscale Walkability for Rolândia

From the sampling of pre-located points distributed every 100 meters on the road network of Rolândia city, using ArcGIS 10.6 software, a total of n=4553 points were selected.

During the image acquisition process using the Google Street View API platform, a total of  $n=3396$  images were obtained formed by  $n=13584$  images that compose the panorama (Figure 14). After processing the obtained images with the semantic segmentation model, the 19 categories of the Cityscapes dataset were used to determine the pixel composition of the sampled images. The absolute pixel count, proportion, and score for each segment, class, and image were calculated. For the statistical analyses conducted in this study, z-scores were employed as the adopted values.

Figure 14- Sampling of geo-tagged points for image extraction in Rolândia.



Source: The author, 2023.

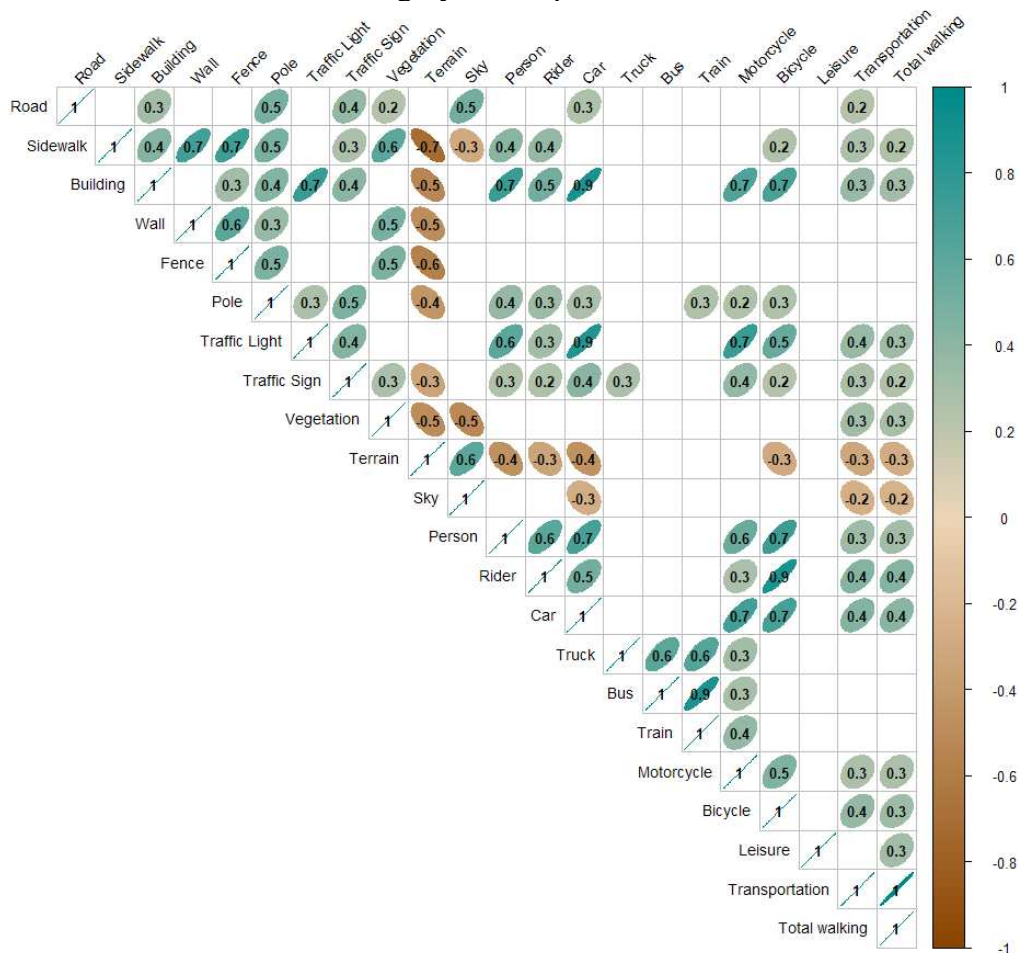
A preliminary statistical analysis was conducted by performing a non-parametric spearman correlation analysis between the absolute number of trips taken per census tracts and the adopted categories for semantic segmentation (Figure 15). This analysis divides walking into utilitarian transportation walking and recreational leisure walking, while considering total walking to examine, in an aggregate way that might inform following analysis, potential differences between walking types and their relationships with micro-scale variables.

In the travel sample obtained from the OD survey, leisure walking did not demonstrate any significant correlation with any micro-scale classes. This outcome could be attributed to an insufficient sample size for analysis, as recreational walking trips only comprised 5.1% ( $n=20$ ) of the total. Conversely, transportation walking showed correlations with specific variables, but

these correlations were neither strong nor very strong. Total walking did exhibit significant correlations, though they were also not particularly strong.

Transportation walking was found to be correlated with the presence of Traffic Lights, the presence of individuals riding motorcycles or bicycles in the Rider category, and the presence of Cars and Bicycles. Total walking showed stronger correlations with only the presence of Cars and Riders. The results align with what one would expect in a city the size of Rolândia. Given the limited sample of leisure walking, both transportation walking and walking in general are likely to be concentrated in areas with extensive urban infrastructure. Such areas, including commercial centers and denser regions, typically have a higher prevalence of mixed-use spaces. Additionally, a negative correlation was observed between transportation walking as well as total walking and the presence of vacant lots and open sky.

Figure 15 - Spearman’s rho correlation between absolute walking trips per census tract and identified category values per zone – Rolândia.

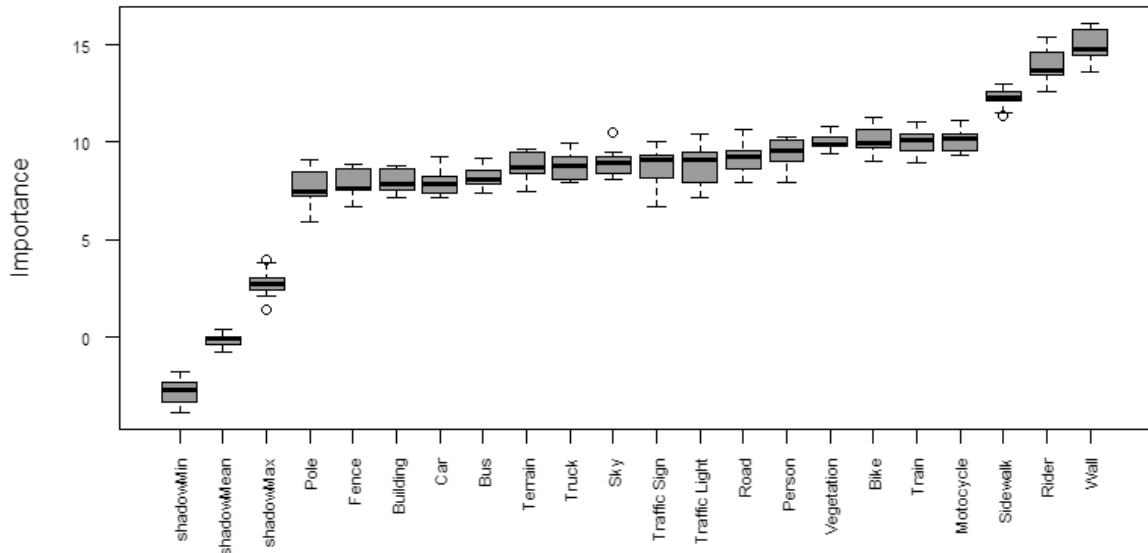


Source: The author, 2023.

The analysis focuses on being predictive based on total walking to verify its relationship with micro-scale variables, moderated by sociodemographic characteristics. Initially, a variable

selection procedure was performed using the Boruta method applied to all categories of semantic segmentation, with the results presented in Figure 16. The order presented in this procedure can be used to obtain the best-fitting model.

Figure 16 - Boxplot representing the importance of micro-scaled categories obtained from semantic segmentation to walking (or not) in Rolândia.



Source: The author, 2023.

The VIF smaller than 9, despite being slightly higher, the obtained values were considered satisfactory. Three models were adjusted for micro-scale variables moderated by demographic variables as predictors of walking in the city of Rolândia. The initial model included environmental variables, the complete model included all micro-scale environmental variables and individual variables, and a better-fitting model was also developed. This improved model indicates that among the individual variables, with children as the reference category for age, adults have a 0.53 OR of walking compared to children, while the elderly have a significantly higher chance of walking, with an OR of 1.75 compared to children. As for gender, women have at least twice the OR of walking compared to men (OR= 0. 498) (Table 10).

In the best-fitted model, the presence of Bicycles was negatively linked to the probability of choosing walking as a mode of transport, with an OR of 0.43. Similarly, an increased presence of Poles, traffic signs for example, in the urban environment corresponds to a decreased OR of 0.34 for walking. However, the presence of Walls and the Rider category (those on motorcycles or bicycles) positively predicts walking. The OR for the 'Rider' category is 8.41, while for the 'walls' category, it stands at 2.41, therefore a higher frequency of these elements corresponds to increased odds of walking.

A spatial concentration analysis was conducted using the Getis-Ord  $G_i^*$  statistic in a Hotspots Coldspots analysis, which identifies statistically significant clusters or hotspots and coldspots with high or low statistical values, respectively (Figure 17). This analysis was performed for all micro-scale variables that were relevant to the best-fitting model. Results for the bicycle presence category show a clear cluster and significant grouping of high values in the central areas of the city, particularly in mixed-use and residential density areas. Regarding the Pole category, significant clusters were identified in the central areas and in the southwest outskirts of the city, where there is a large cluster of service and industrial usage. In relation to the category of individuals riding motorcycles or bicycles, a cluster is also observed in the city center, along with a cluster of negative values in peripheral areas. As for the walls micro-scale feature, concentration areas of higher values are mainly found outside the city's commercial center, in residential areas below the historic center, predominantly residential areas, and industrial areas of the city.

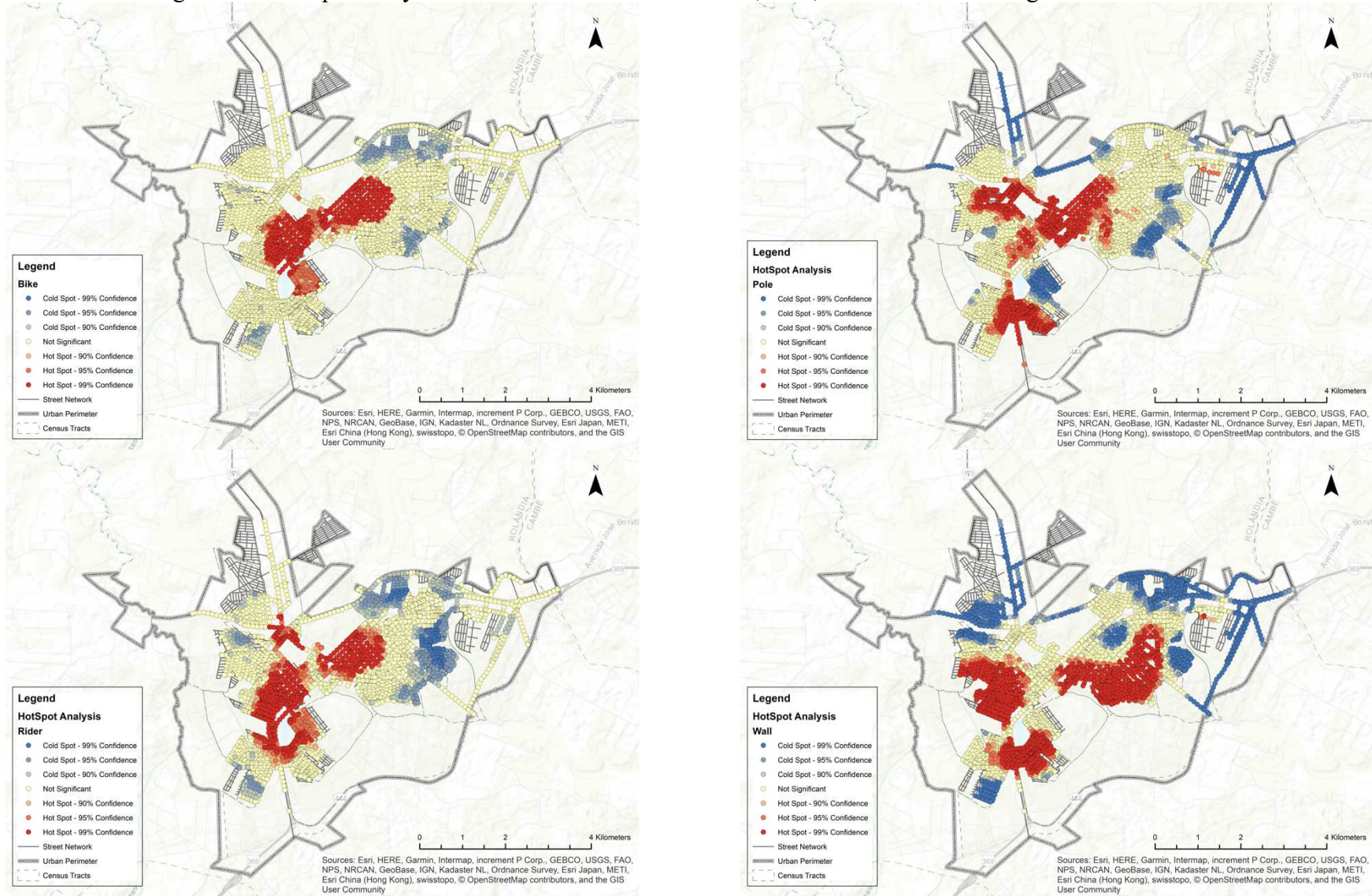
Table 10 - Multilevel logistic regression analysis on walking, individual factors and micro-level-built environment characteristics for Rolândia (n=2152).

| Model Predictor  | (1) Micro BE variables                   |                 |          | (2) Complete model                     |                 |          | (3) Best fitted model                       |                 |          |
|--|--|-----------------|----------|--|-----------------|----------|---|-----------------|----------|
|  | OR                                       | 95% CI          | p-value  | OR                                     | 95% CI          | p-value  | OR  | 95% CI          | p-value  |
| <b>Individual</b>  |  |                 |          |  |                 |          |   |                 |          |
| <b>Age</b>   |  |                 |          |  |                 |          |   |                 |          |
| Child(≤12)   | -  | -               | -        | 1.00*                                  | -               | -        | 1.00*                                       | -               | -        |
| Adult (<13 to 59<)   | -  | -               | -        | 0.542                                  | 0.371<br>0.791  | <0.001** | 0.534                                       | 0.368<br>0.776  | <0.001** |
| Elderly (≤60)  | -  | -               | -        | 1.788                                  | 0.108<br>0.293  | <0.05*   | 1.751                                       | 1.075<br>2.854  | <0.05*   |
| <b>Gender</b>  |  |                 |          |  |                 |          |   |                 |          |
| Female   | -  | -               | -        | 1.00*                                  | -               | -        | 1.00*                                       | -               | -        |
| Male   | -  | -               | -        | 0.500                                  | 0.385<br>0.649  | <0.001** | 0.498                                       | 0.384<br>0.645  | <0.001** |
| Absent   | -  | -               | -        | 0.084                                  | 0.191<br>0.371  | <0.001** | 0.082                                       | 0.018<br>0.371  | <0.001** |
| <b>Household income</b>  |  |                 |          |  |                 |          |   |                 |          |
| ≤ 2 MW - Reference   | -  | -               | -        | 1.00*                                  | -               | -        | 1.00*                                       | -               | -        |
| <2 to 5 MW<  | -  | -               | -        | 0.356                                  | 0.138<br>0.914  | <0.05*   | 0.351                                       | 0.135<br>0.909  | 0.05*    |
| ≤ 5 MW   | -  | -               | -        | 0.000                                  | 0.000<br>0.001  | 0.957    | 0.000                                       | 0.000<br>0.000  | 0.98530  |
| Absent   | -  | -               | -        | 2.127                                  | 1.594<br>2.839  | <0.001** | 2.093                                       | 1.571<br>2.789  | <0.001** |
| <b>Built environment (cont.var/ average z-scores per zone)</b> |  |                 |          |  |                 |          |   |                 |          |
| Bike   | 0.266                                    | 0.134<br>0.213  | <0.05*   | 0.198                                  | 0.053<br>0.743  | <0.05*   | 0.438                                       | 0.163<br>1.179  | 0.102    |
| Building   | 1.891                                    | 0.067<br>1.056  | 0.340    | 2.965                                  | 0.839<br>10.469 | 0.091 .  | -   | -               | -        |
| Bus  | 0.069                                    | 0.510<br>7.003  | 0.073    | 0.078                                  | 0.004<br>1.381  | 0.081 .  | -   | -               | -        |
| Car  | 0.896                                    | 0.003<br>1.289  | 0.876    | 0.663                                  | 0.175<br>2.506  | 0.545    | -   | -               | -        |
| Fence  | 0.451                                    | 0.227<br>3.538  | 0.133    | 0.427                                  | 0.154<br>1.185  | 0.102    | -   | -               | -        |
| Motorcy.   | 0.627                                    | 0.159<br>1.277  | 0.607    | 0.623                                  | 0.113<br>3.426  | 0.587    | -   | -               | -        |
| Person   | 1.833                                    | 0.105<br>3.719  | 0.294    | 1.627                                  | 0.542<br>4.876  | 0.384    | -   | -               | -        |
| Pole   | 0.284                                    | 0.589<br>5.695  | <0.001** | 0.339                                  | 0.131<br>0.872  | <0.05*   | 0.342                                       | 0.158<br>0.741  | <0.001** |
| Rider  | 6.861                                    | 0.107<br>0.753  | <0.001** | 9.446                                  | 2.281<br>39.109 | <0.001** | 8.415                                       | 2.294<br>30.862 | <0.001** |
| Road   | 2.089                                    | 1.591<br>29.585 | 0.170    | 2.062                                  | 0.728<br>5.840  | 0.172    | -   | -               | -        |
| Sidewalk   | 1.507                                    | 0.727<br>6.000  | 0.411    | 1.306                                  | 0.493<br>3.459  | 0.590    | -   | -               | -        |
| Sky  | 0.797                                    | 0.566<br>4.010  | 0.799    | 0.695                                  | 0.118<br>4.070  | 0.686    | -   | -               | -        |
| Terrain  | 1.349                                    | 0.138<br>4.588  | 0.668    | 1.467                                  | 0.377<br>5.709  | 0.579    | -   | -               | -        |
| Train  | 0.583                                    | 0.342<br>5.309  | 0.627    | 0.509                                  | 0.062<br>4.140  | 0.528    | -   | -               | -        |
| Traffic Light  | 0.582                                    | 0.066<br>5.153  | 0.392    | 0.459                                  | 0.136<br>1.543  | 0.208    | -   | -               | -        |
| Traffic Sign   | 0.455                                    | 0.167<br>2.014  | 0.487    | 0.544                                  | 0.061<br>4.825  | 0.585    | -   | -               | -        |
| Truck  | 1.811                                    | 0.049<br>4.214  | 0.407    | 1.985                                  | 0.508<br>7.748  | 0.323    | -   | -               | -        |
| Vegetation   | 1.057                                    | 0.444<br>7.369  | 0.953    | 1.104                                  | 0.168<br>7.228  | 0.917    | -   | -               | -        |
| Wall   | 3.210                                    | 0.164<br>6.785  | <0.05*   | 2.579                                  | 1.117<br>5.951  | <0.05*   | 2.414                                       | 1.399<br>4.167  | <0.001** |
| <b>Model Fit</b>   | <b>AIC = 1917.9,<br/>logLik = -937.9</b> |                 |          | <b>AIC = 1799.6<br/>logLik= -871.8</b> |                 |          | <b>AIC = 1782.792***,<br/>logLik=-878.4</b> |                 |          |

Note!: \*=1,00: reference category; CI 95%: 95% Confidence Interval. \*\* significative at the  $\alpha=0,05$  level. \*\*\* ANOVA's  $\Pr(>Chisq)$  at  $\alpha=0.05$

Source: The author. 2023.

Figure 17 – Hotspot analysis for the distribution of the Bike, Pole, Rider and Wall categories for Rolândia.



Source: The author, 2023.

## 4.2 Results for Cambé

Table 11 presents the sociodemographic characteristics of the sample population who participated in the origin destination survey conducted in Cambé, comprising a total of 3348 individuals. The table encompasses two main variables: household income and walking behavior. Regarding household income, the sample is divided into two categories. The first category, representing 78.67% of the total sample, consists of 2634 individuals with a household income equal to or less than 2 MW. The second category comprises 714 individuals, accounting for 21.33% of the sample, whose household income falls within the range of more than 2 MW but less than 5 MW.

The walking behavior of the participants is also analyzed in the table. Among the sample population, 949 individuals, constituting 28.34% of the total, reported engaging in walking. On the other hand, 2399 individuals, making up 71.65% of the sample, indicated that they do not engage in walking.

Table 11 - Sociodemographic characteristics of the sample for Cambé (n =3348).

| Sociodemographic Variables     | Count | Percentage (%) |
|--------------------------------|-------|----------------|
| <b><i>Household income</i></b> |       |                |
| ≤ 2 MW                         | 2634  | 78.67          |
| <2 to 5 MW<                    | 714   | 21.33          |
| <b><i>Walking</i></b>          |       |                |
| Yes                            | 949   | 28.34          |
| No                             | 2399  | 71.65          |

Source: The author, 2023.

Table 12 displays the modal response for each sociodemographic variable, distinguishing between individuals who reported walking and those who provided other responses. The sociodemographic variable analyzed in this table is household income based on data from the national Brazilian census from 2010 (IBGE, 2010). For individuals who reported walking, the majority (785) had a household income of ≤ 2 MW, while a smaller proportion (164) fell within the income range of <2 to 5 MW. On the other hand, individuals who provided responses other than walking showed different patterns. Among this group, 1849 individuals had a household income of ≤ 2 MW, while 550 individuals fell within the income range of <2 to 5 MW.

Table 12 - Sociodemographic characteristics of the sample versus walking or not in Cambé (n = 3348).

| Sociodemographic Variables   | Modal |         |
|------------------------------|-------|---------|
|                              | Other | Walking |
| Household income $\leq 2$ MW | 1849  | 785     |
| $<2$ to $5$ MW $<$           | 550   | 164     |

Source: The authors. 2023.

Figure 18 presents a choropleth map that visually depicts the aggregate absolute number of walking trips conducted by census tracts in Cambé. This figure serves to illustrate the distribution and intensity of walking trips throughout the city, utilizing data collected from the origin destination survey. The color scheme employed in the map effectively represents the varying levels of absolute walking trip numbers, with distinct shades denoting higher or lower values. Notably, the census tract situated in the downtown area, known for its substantial presence of commerce and services, exhibits the highest value on the map. This observation indicates a concentrated prevalence of walking trips within this central region. Furthermore, apart from the downtown area, other census tracts displaying elevated values predominantly exist in the city center and the southeastern region, which experiences urban agglomeration (conurbation) with the neighboring city of Londrina. These specific areas are characterized by a dense concentration of residences and localized commercial services, consequently leading to an increased volume of walking trips.

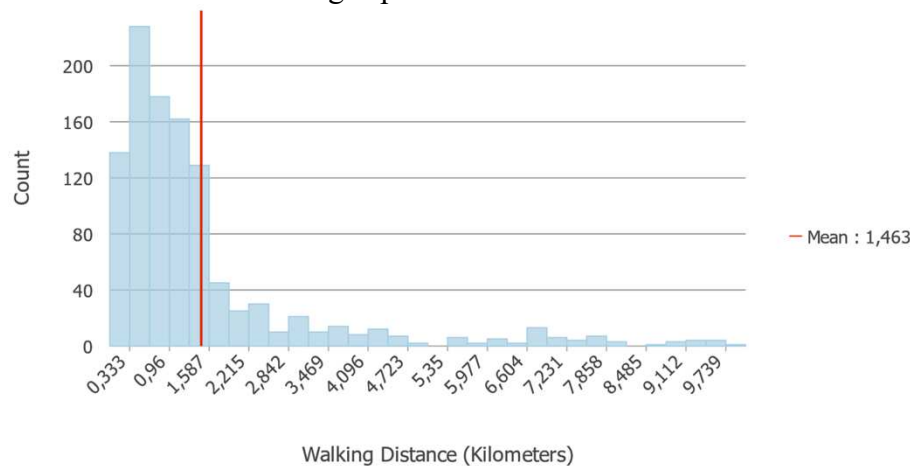
Figure 18 - Representation of the aggregate absolute number of walking trips conducted by census tracts in Cambé.



Source: The authors. 2023.

Figure 19 depicts a histogram representing the walking distances of the n=1080 walking trips conducted in Cambé. The histogram provides a visual representation of the distribution of walking distances within the dataset. The x-axis of the histogram represents the walking distances in kilometers, while the y-axis displays the frequency or count of walking trips falling within each distance interval. The highest concentration of values is observed at approximately 0.96 kilometers, indicating that a significant number of walking trips have a distance close to this value. The mean walking distance, calculated as 1.463 kilometers, provides an estimate of the average distance covered by individuals in the sample during their walking trips. This value can serve as a measure of central tendency, representing the typical distance traveled by pedestrians in Cambé.

Figure 19 – Histogram of walking distances of the n=1080 walking trips conducted in Cambé.



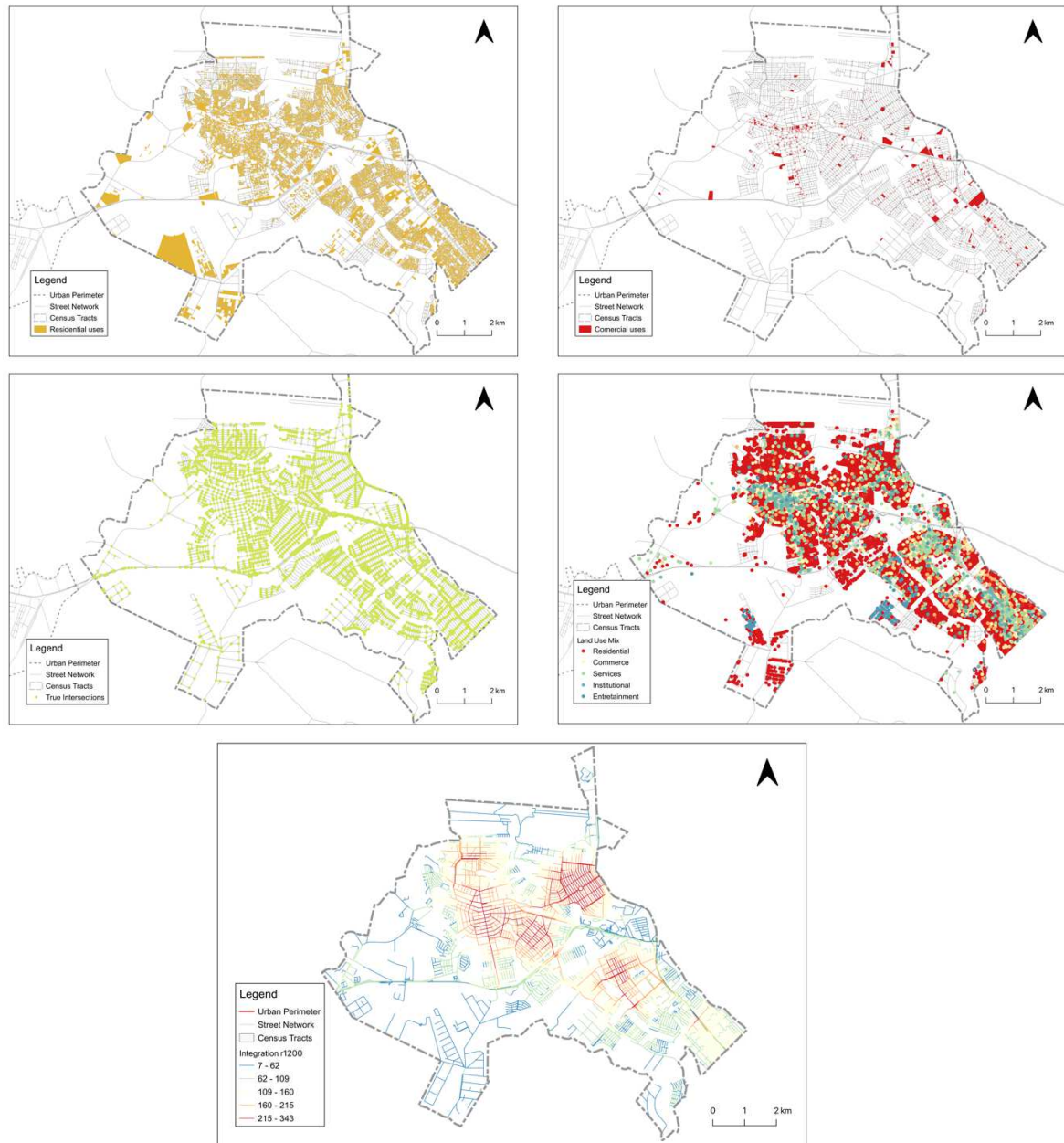
Source: The authors. 2023. Based on data provided by Cambé's administration (2019).

#### 4.2.1 Macroscale Walkability for Cambé

Figure 20 presents five distinct variables used for measuring walkability macro scale in Cambé. These data include information on residential parcels, built area for commercial purposes, intersections, land use data, and street integration calculations from space syntax. The dataset on residential parcels provides details on the spatial distribution and boundaries of residential areas in Cambé, facilitating analysis of residential land use patterns and spatial organization. The built area dataset focuses on the physical extent of buildings used for commercial activities in the city. It enables assessment of the spatial distribution and density of commercial structures. The dataset on intersections encompasses the locations and characteristics of intersections within Cambé's road network. It allows for analysis of street connectivity, traffic flow, and transportation network efficiency. The land use dataset provides

information on the various types of land use in Cambé, including residential, commercial, industrial, and recreational areas. It aids in understanding the distribution and composition of land uses across the city. Lastly, the street integration calculations from space syntax applies analysis techniques to evaluate street integration measures for Cambé's street network.

Figure 20 - Data used for measuring variables at the macro scale for Cambé. Respectively: residential parcels, built area (building footprint) for commercial purposes, intersections, land use data, street integration calculations from space syntax.



Source: The author, 2023.

The multilevel logistic regression analysis conducted for micro-scale variables in the city of Cambé revealed an ICC for the null model of 0.21, indicating that around 21% of the variability in the outcome variable can be attributed to variations between different zones within

the city. A VIF value smaller than 2 was observed in the analysis of the best fitted model presented in Table 13. This is probably due to the small number of elements in the model.

The results of a multilevel logistic regression analysis exploring the relationship between the choice of walking, micro/macro built environment factors, and individual factors in Cambé are presented on Table 13. The study sample consists of 3348 individuals. The table provides results for three models: the first includes only individual factors, the second represents the complete model, and the third presents the best fitted model.

Regarding individual factors, household income is examined, with  $\leq 2$  MW as the reference category. The results show that individuals with a household income ranging from  $<2$  to 5 MW have a non-significant association with the choice of walking compared to the reference category. The table also includes measures such as land use mix (entropy), residential density, intersection density, retail FAR, and integration r1200. None of these variables exhibit a significant association with the choice of walking in the complete or best fitted models. The model fit statistics, such as AIC and log-likelihood, are reported for each model to assess their goodness of fit. The AIC values indicate that the complete model and the best fitted model have similar fit to the data, with the latter showing a slightly lower AIC value. Within a comparison using the new AICs of the models, no significant differences were identified between a complete model with environmental and individual variables, a model with only environmental variables, and no significant differences have been found between a model with no combination of environmental variables and the sociodemographic variable on household income.

In the analysis, it was observed that no significant associations were found between the choice of walking and individual factors or the evaluated built environment macro variables in Cambé. This absence of significance might be due to Cambé's distinct context, potentially shaped by urban agglomeration effects or other factors not considered in the analysis. The socioeconomic variable derived from the IBGE 2010 census data did not effectively mediate the relationship between the studied variables. One hypothesis is that this result stemmed from outdated income data from the IBGE or its aggregation, or possibly the mismatch in temporal alignment between the city's urban environment data and these socioeconomic data. It's essential to underscore that this is cross-sectional research, and the findings provide a snapshot from 2023, limited by the data sources available at the time.

Table 13 - Multilevel logistic regression analysis for the relationship between the choice of walking, macro/macro-built environment, and individual factors for Cambé (n = 3348).

| Model Predictor  | (1) BE variables                         |                |         | (2) Complete model                       |                |         | (3) Best fitted model |        |         |
|--|--|----------------|---------|--|----------------|---------|-----------------------|--------|---------|
|  | OR                                       | 95% CI         | p-value | OR                                       | 95% CI         | p-value | OR                    | 95% CI | p-value |
| <b>Individual</b>  |  |                |         |  |                |         |                       |        |         |
| <b>Household income</b>  |  |                |         |  |                |         |                       |        |         |
| ≤ 2 MW - Reference   | -  | -              | -       | 1.00*                                    | -              | -       | -                     | -      | -       |
| <2 to 5 MW<  | -  | -              | -       | 0.805                                    | 0.528<br>1.229 | 0.3167  | -                     | -      | -       |
| <b>Built environment (cont.var/ average z-scores per zone)</b> |  |                |         |  |                |         |                       |        |         |
| Land use Mix (entropy)   | 0.995                                    | 0.862<br>1.147 | 0.939   | 1.010                                    | 0.874<br>1.167 | 0.8867  | -                     | -      | -       |
| Residential Density  | 1.118                                    | 0.953<br>1.312 | 0.172   | 1.087                                    | 0.920<br>1.285 | 0.3221  | -                     | -      | -       |
| Intersection Density   | 0.991                                    | 0.864<br>1.136 | 0.893   | 1.000                                    | 0.873<br>1.147 | 0.9920  | -                     | -      | -       |
| Retail FAR   | 1.043                                    | 0.920<br>1.183 | 0.509   | 1.042                                    | 0.920<br>1.180 | 0.5091  | -                     | -      | -       |
| Integration r1200  | 0.951                                    | 0.469<br>1.930 | 0.890   | 1.063                                    | 0.528<br>1.229 | 0.8687  | -                     | -      | -       |
| <b>Model Fit</b>   | <b>AIC = 3948.5<br/>logLik = -1967.3</b> |                |         | <b>AIC = 3949.5<br/>logLik = -1966.8</b> |                |         | -                     |        |         |

Note!: \*=1,00: reference category; CI 95%: 95% Confidence Interval. \*\* significative at the  $\alpha=0,05$  level.

Source: The author, 2023.

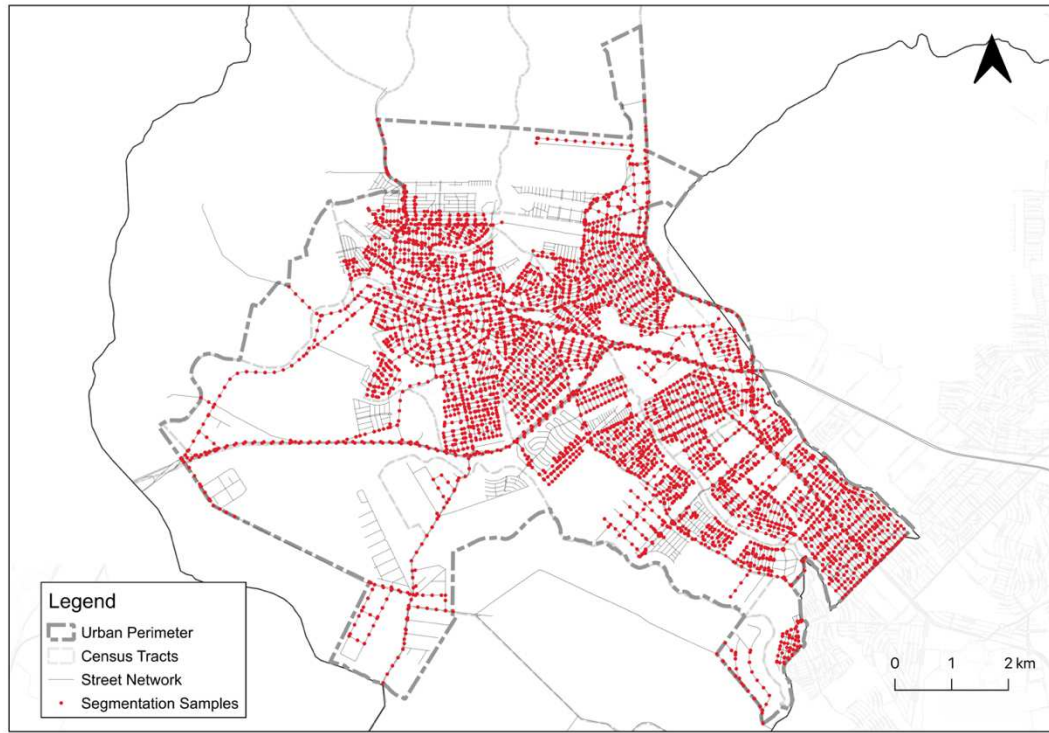
#### 4.2.2 Microscale Walkability for Cambé

To evaluate the walkability at the microscale level in Cambé, a sampling approach was implemented, involving the extraction of geo-tagged images. Figure 21 displays the sampling strategy used to select these images.

Using the ArcGIS 10.6 software, pre-located points were distributed at regular intervals of 100 meters along the road network of Cambé. These points served as the basis for capturing images using the Google Street View API platform. As a result, a total of n=4487 images were successfully obtained for further analysis composed by n=17948 images that form the panorama. The acquired images were subjected to processing utilizing a semantic segmentation model. This model utilized the 19 categories available in the Cityscapes dataset to determine the pixel composition within each sampled image. Through this analysis, the absolute pixel count, proportion, and score were calculated for each segment, class, and image. The statistical analyses conducted in this study, z-scores were utilized as the adopted values. These z-scores

allowed for the standardization of the obtained results, enabling comparisons and assessments of the walkability features at the microscale level in Cambé.

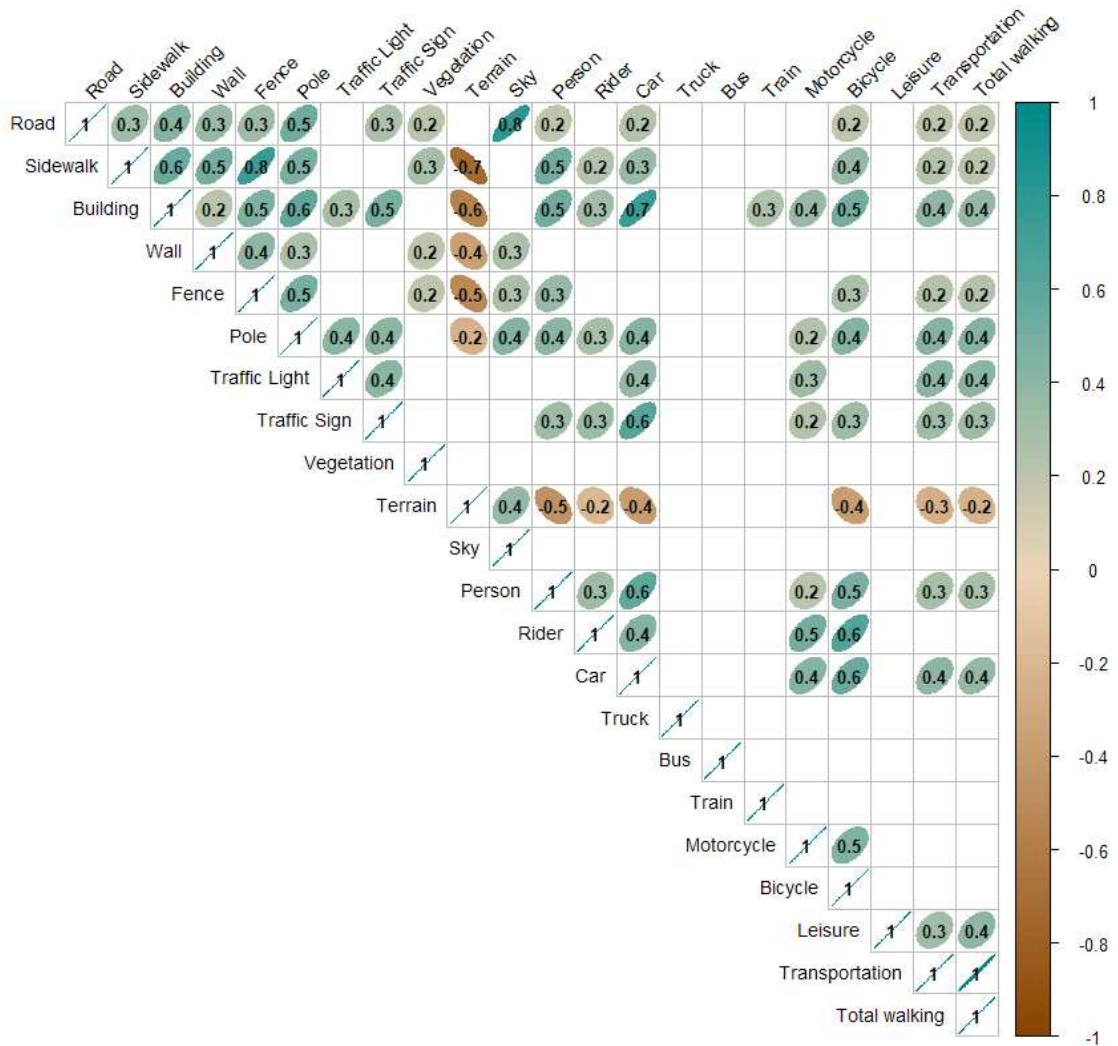
Figure 21 - Sampling of geo-tagged points for image extraction in Cambé.



Source: The author, 2023.

An initial analysis involved the examination of Spearman correlations, which is a non-parametric correlation analysis, between microscale variables and total walking, walking for transportation, and walking for leisure in the city of Cambé (Figure 22). It was observed that walking for leisure did not exhibit any significant correlations with the microscale variables. On the other hand, walking for transportation and total walking displayed some relevant but not very strong correlations. The highest correlations for both walking for transportation and total walking were found with the presence of street art, the presence of streetlights or traffic lights, and the presence of vehicles. This indicates that higher values of these microscale categories are associated with greater levels of walking within the sector, whether it be for transportation or overall walking.

Figure 22 - Spearman’s rho correlation between absolute walking trips per zone and identified category values per zone – Cambé

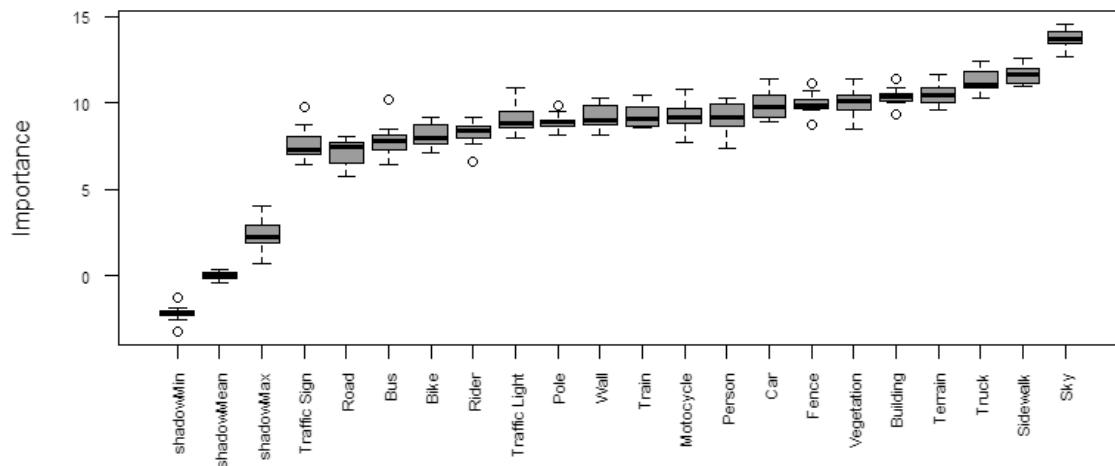


Source: The author, 2023.

Figure 23 presents a boxplot depicting the importance of micro-scaled categories derived from semantic segmentation in relation to walking behavior in Cambé. The importance of these categories was determined using the Boruta feature selection method, which aids in identifying the most influential variables. The order obtained from this selection procedure was subsequently employed to introduce the variables into the regression modeling for testing the best fit. The boxplot illustrates the distribution and variability of the importance scores assigned to each micro-scaled category. The importance scores reflect the degree to which a particular category contributes to the prediction of walking behavior. A higher importance score suggests a stronger association between the category and walking behavior. The boxplot provides a visual representation of the range of importance scores and any potential outliers among the micro-scaled categories. It allows for a comparative assessment of the relative importance of

each category in predicting walking behavior. The utilization of the Boruta feature selection method ensures that only the most relevant and significant micro-scaled categories are included in the regression modeling process. By employing the order derived from this method, the regression model can be built based on the variables that offer the best fit and predictive power.

Figure 23 - Boxplot representing the importance of micro-scaled categories obtained from semantic segmentation to walking (or not) in Cambé.



Source: The author, 2023.

Table 14 presents the results of a multilevel logistic regression analysis examining the relationship between walking behavior, individual factors, and micro-level built environment characteristics in Cambé. The study sample consists of  $n=3348$  trips conducted by any modal. The table includes three models: the first model consists of individual factors, the second model represents the complete model, and the third model shows the best fitted model. The best fitted model presented a VIF value below 2, indication of low multicollinearity. Regarding individual factors, household income is considered, with  $\leq 2$  MW as the reference category. The results indicate that individuals with a household income between  $<2$  to 5 MW do not exhibit a significant correlation with walking behavior in both the complete and best fitted models.

The model fit statistics, such as AIC and log-likelihood, are reported for each model, indicating their goodness of fit. The best fitted model shows the lowest AIC value, suggesting a better fit to the data. In conclusion, the multilevel logistic regression analysis conducted in Cambé with walking as a binary response and micro-scale BE features and individual factors reveals that household income measured utilizing the IBGE census information, does not significantly correlate with walking behavior. Examining the micro-level-built environment variables, it is observed that most of these do not have a significant correlation with walking

behavior in either the complete or best fitted models. However, the presence of Terrain and Vegetation demonstrates some level of significance in explaining walking behavior in the best fitted model, with lower ORs and wider confidence intervals. These findings highlight the complex relationship between walking behavior and individual as well as environmental factors in Cambé.

The best-fitted model for micro-scale variables in the city of Cambé revealed that the Terrain and Vegetation exhibited statistically significant relationships in explaining walking. The ORs for walking decreased as the values of these variables decreased, with OR of 0.60 and 0.64, respectively. Furthermore, the spatial analysis of hotspots (Figures 24 and 25) highlighted certain patterns in the distribution of these variables. Concentrations of vacant lots (Terrain) were predominantly observed in the peripheral areas surrounding the city center, while co-occurring spots were identified in densely populated and urbanized regions, particularly in proximity to the city of Londrina. These findings suggest that areas characterized by a high density of vegetation and vacant lots are negatively associated with walking behavior. In simpler terms, the presence of such characteristics reduces the likelihood of individuals engaging in walking trips, especially in peripheral areas. It is essential to emphasize that this result does not necessarily reflect people's preferences for walking locations but rather indicates the actual areas where walking occurs.

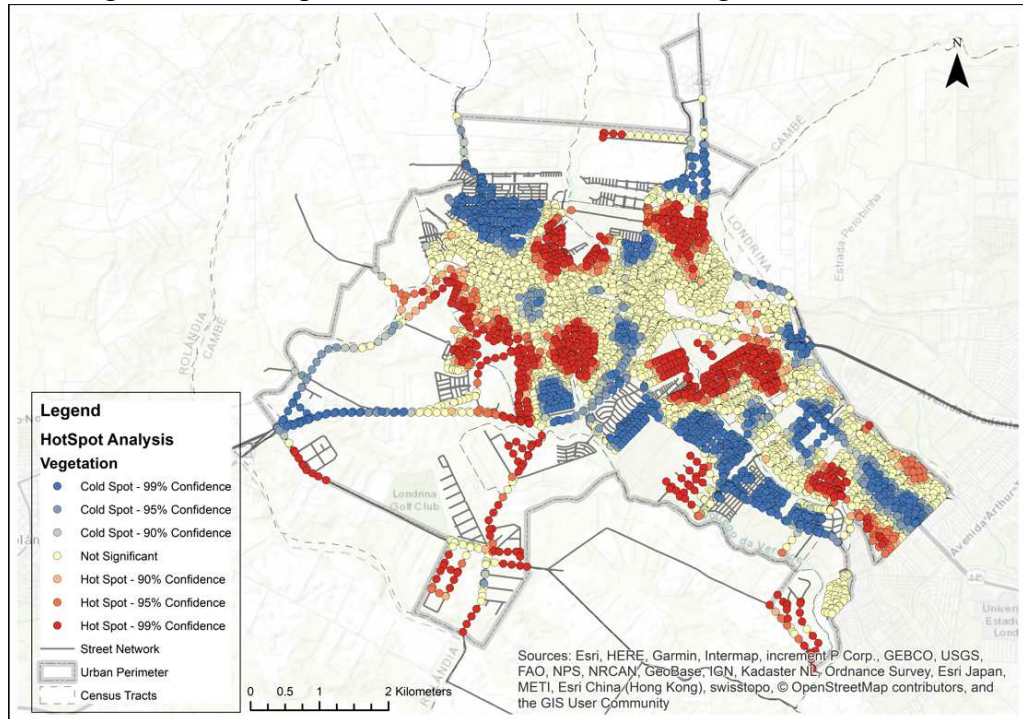
Table 14 - Multilevel logistic regression analysis for the relationship between walking or not, individual factors and micro-level-built environment characteristics for Cambé (n=3348).

| Model Predictor  | (1) Micro BE variables                    |                |         | (2) Complete model                      |                |          | (3) Best fitted model                      |                |         |
|--|---|----------------|---------|---|----------------|----------|--|----------------|---------|
|  | OR  | 95% CI         | p-value | OR                                      | 95% CI         | p-value  | OR   | 95% CI         | p-value |
| <b>Individual</b>  |   |                |         |   |                |          |  |                |         |
| <b>Household income</b>  |   |                |         |   |                |          |  |                |         |
| ≤ 2 MW - Reference   | -   | -              | -       | 1.00*                                   | -              | -        | 1.00*                                      | -              | -       |
| <2 to 5 MW<  | -   | -              | -       | 0.742                                   | 0.472<br>1.167 | 0.197    | 0.69562                                    | 0.479<br>1.009 | 0.005*  |
| <b>Built environment (cont.var/ average z-scores per zone)</b> |   |                |         |   |                |          |  |                |         |
| Bike   | 1.201                                     | 0.580<br>2.487 | 0.694   | 1.053                                   | 0.493<br>2.246 | 0.8937   | -  | -              | -       |
| Building   | 0.355                                     | 0.111<br>1.133 | 0.616   | 0.425                                   | 0.130<br>1.390 | 0.1572   | -  | -              | -       |
| Bus  | 0.687                                     | 0.340<br>1.388 | 0.581   | 0.736                                   | 0.359<br>1.507 | 0.4024   | -  | -              | -       |
| Car  | 0.727                                     | 0.369<br>1.433 | 0.619   | 0.813                                   | 0.405<br>1.632 | 0.5620   | -  | -              | -       |
| Fence  | 0.834                                     | 0.393<br>1.770 | 0.621   | 0.874                                   | 0.411<br>1.857 | 0.7273   | -  | -              | -       |
| Motorcycle   | 0.903                                     | 0.485<br>1.682 | 0.601   | 0.832                                   | 0.445<br>1.558 | 0.5676   | -  | -              | -       |
| Person   | 1.304                                     | 0.758<br>2.243 | 0.672   | 1.267                                   | 0.739<br>2.171 | 0.3888   | -  | -              | -       |
| Pole   | 0.659                                     | 0.367<br>1.184 | 0.600   | 0.655                                   | 0.365<br>1.174 | 0.1558   | -  | -              | -       |
| Rider  | 0.972                                     | 0.470<br>2.013 | 0.630   | 1.008                                   | 0.489<br>2.078 | 0.9817   | -  | -              | -       |
| Road   | 0.340                                     | 0.088<br>1.310 | 0.618   | 0.368                                   | 0.095<br>1.420 | 0.1470   | -  | -              | -       |
| Sidewalk   | 0.447                                     | 0.141<br>1.421 | 0.617   | 0.474                                   | 0.149<br>1.505 | 0.2057   | -  | -              | -       |
| Sky  | 0.191                                     | 0.022<br>1.624 | 0.617   | 0.210                                   | 0.025<br>1.774 | 0.1520   | -  | -              | -       |
| Terrain  | 0.182                                     | 0.031<br>1.052 | 0.617   | 0.215                                   | 0.036<br>1.269 | 0.0897   | 0.606                                      | 0.432<br>0.850 | 0.05*   |
| Train  | 1.224                                     | 0.545<br>2.749 | 0.568   | 1.345                                   | 0.599<br>3.021 | 0.4717   | -  | -              | -       |
| Traffic Light  | 0.794                                     | 0.388<br>1.625 | 0.666   | 0.792                                   | 0.390<br>1.608 | 0.5199   | -  | -              | -       |
| Traffic Sign   | 1.327                                     | 0.733<br>2.403 | 0.608   | 1.325                                   | 0.730<br>2.405 | 0.3536   | -  | -              | -       |
| Truck  | 0.470                                     | 0.228<br>0.971 | 0.616   | 0.440                                   | 0.214<br>0.906 | 0.0260 * | -  | -              | -       |
| Vegetation   | 0.091                                     | 0.007<br>1.084 | 0.623   | 0.104                                   | 0.008<br>1.234 | 0.0730.  | 0.646                                      | 0.442<br>0.944 | 0.05*   |
| Wall   | 1.201                                     | 0.580<br>2.487 | 0.694   | 1.053                                   | 0.493<br>2.246 | 0.8937   | -  | -              | -       |
| <b>Model Fit</b>   | <b>AIC = 3956.5,<br/>logLik = -1958.3</b> |                |         | <b>AIC = 3956.5<br/>logLik= -1957.2</b> |                |          | <b>AIC = 3934.3***,<br/>logLik=-1962.1</b> |                |         |

Note!: \*=1,00: reference category; CI 95%: 95% Confidence Interval. \*\* significative at the  $\alpha=0,05$  level.  
\*\*\* ANOVA's  $\text{Pr}( > \text{Chisq} )$  at  $\alpha=0.05$

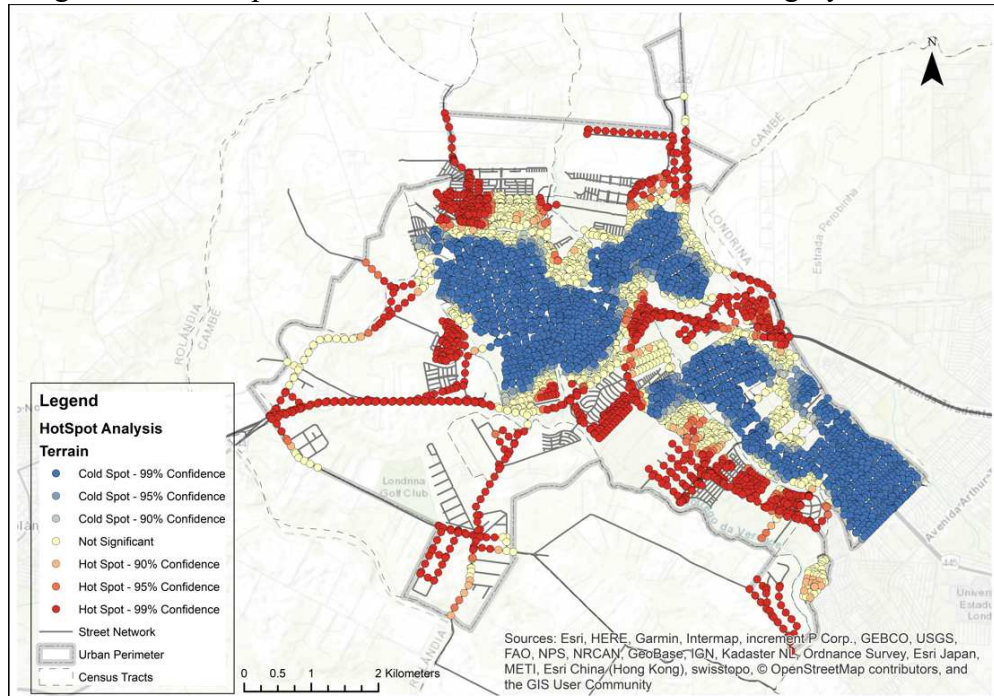
Source: The author, 2023.

Figure 24 - Hot Spot for the distribution of the Vegetation in Cambé.



Source: The author, 2023.

Figure 25 - Hot Spot for the distribution of the Terrain category in Cambé.



Source: The author, 2023.

### 4.3 Results for Londrina

In Londrina-PR the n=18635 valid trips conducted by any modal extracted from the municipal OD survey were considered the study sample. From the 18635 from the sample 4162 were walking trips (22.33%), This is the general response variable for the statistical analyses to be conducted. The descriptive statistics of the sociodemographic characteristics of the individuals who reported the trips are shown in Table 15. When it comes to gender, the sample is reasonably balanced. However, for *Age*, the 'adult' category is heavily oversampled, which might affect results. Income data is relatively balanced, still being suitable for analysis.

Table 15 - Sociodemographic characteristics of the sample. (n = 18635).

| Sociodemographic Variables     | Count | Percentage (%) |
|--------------------------------|-------|----------------|
| <b><i>Age</i></b>              |       |                |
| Child ( $\leq 12$ )            | 2210  | 11.9           |
| Adult ( $< 13$ to $59 <$ )     | 13496 | 72.4           |
| Elderly ( $\leq 60$ )          | 2929  | 15.7           |
| <b><i>Gender</i></b>           |       |                |
| Female                         | 9374  | 50.3           |
| Male                           | 9261  | 49.7           |
| <b><i>Household income</i></b> |       |                |
| $\leq 2$ MW                    | 4651  | 25.0           |
| $< 2$ to $5$ MW $<$            | 7182  | 38.5           |
| $\leq 5$ MW                    | 2635  | 14.1           |
| Absent                         | 4167  | 22.4           |

Source: The author, 2023.

Table 16 Table 16 presents the distribution of sociodemographic variables and their corresponding walking rates in the study population. The variables include Child ( $\leq 12$ ), Adult ( $< 13$  to  $59 <$ ), Elderly ( $\leq 60$ ), Female, Male,  $\leq 2$  MW (Monthly Wage),  $< 2$  to  $5$  MW $<$ ,  $\leq 5$  MW, and Absent. For the Child category (n=2210), 59.86% of the respondents were in the modal group, while 40.13% fell into the "Other" category. Among Adults (n=13469), the majority, 81.72%, belonged to the modal group, while 18.47% were categorized as "Other". Similarly, the majority of the Elderly (n=2929) were in the modal group, accounting for 73.13% of the population, while 26.86% fell into the "Other" category. Analyzing the data by gender, for Females (n=9374), 74.11% were in the modal group, while 25.89% were classified as "Other". Among Males (n=9261), the majority, 81.26%, belonged to the modal group, while 18.734% were in the "Other" category. When considering income levels,  $\leq 2$  MW (n=4651) had 64.26% of respondents in the modal group, while 35.73% were classified as "Other". For the  $< 2$  to  $5$  MW $<$  category (n=7182), 80.59% were in the modal group, and 19.41% were categorized as

"Other". Among those with  $\leq 5$  MW (n=2635), 87.09% were in the modal group, and 12.90% were classified as "Other". Finally, for the Absent category (n=4167), 81.61% of respondents were in the modal group, and 18.38% were categorized as "Other".

The analysis of sociodemographic variables and their corresponding walking rates in the study population reveals distinct patterns that can be attributed to several influencing factors. Age emerges as a significant determinant, with younger individuals, including both Children and Adults, exhibiting higher rates of walking compared to the Elderly. This discrepancy is likely influenced by factors such as elevated physical activity levels, greater overall mobility, and fewer limitations in mobility often associated with younger age groups. Gender also plays a role, as both males and females in the modal group exhibit higher walking rates than those categorized as "Other." This divergence may stem from cultural and societal influences, as well as variations in activity patterns and mobility preferences between genders. Moreover, income levels exhibit an association with walking rates, whereby the modal group in each income category demonstrates higher rates of walking compared to the "Other" category. This discrepancy could be linked to individuals with higher incomes having greater access to diverse transportation options, thus reducing their reliance on walking. Conversely, individuals with lower incomes may encounter limited access to alternative modes of transportation, necessitating a more frequent reliance on walking for commuting and daily activities.

Table 16- Sociodemographic characteristics of the sample versus walking or not in Londrina (n = 18635).

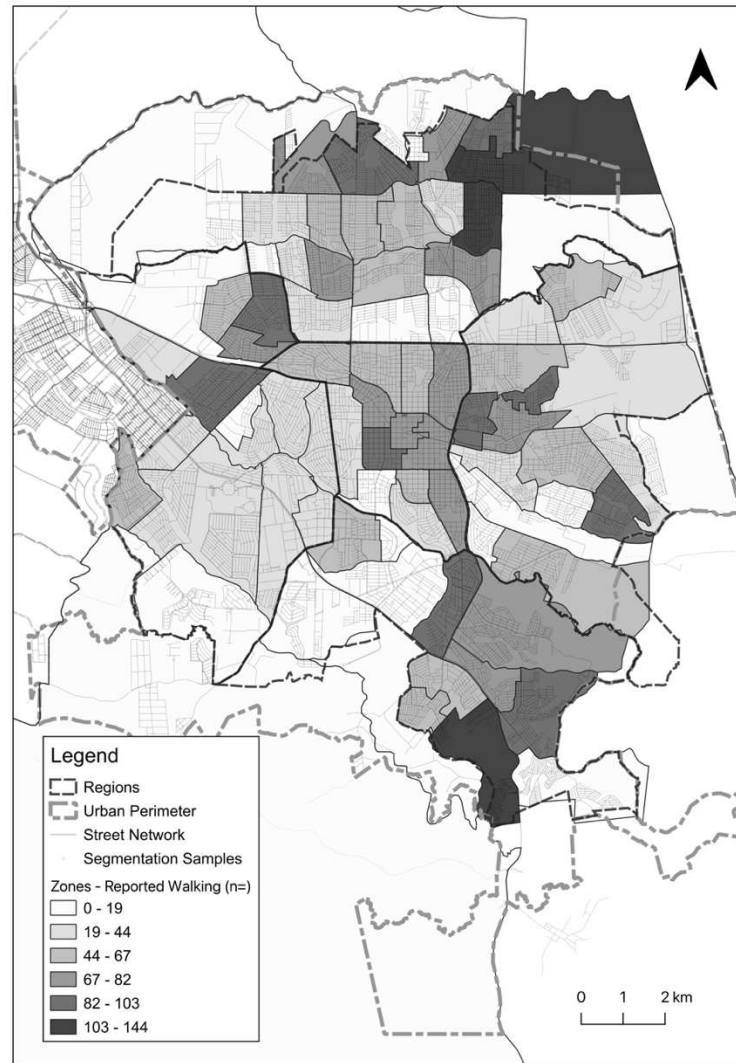
| Sociodemographic Variables     | Modal |       |      |        |
|--------------------------------|-------|-------|------|--------|
|                                | Other | %     | Walk | %      |
| Child ( $\leq 12$ ) (n=2210)   | 1323  | 59.86 | 887  | 40.13  |
| Adult (<13 to 59<) (n=13469)   | 11008 | 81.72 | 2488 | 18.47  |
| Elderly ( $\leq 60$ ) (n=2929) | 2142  | 73.13 | 787  | 26.86  |
| Female (n=9374)                | 6947  | 74.11 | 2427 | 25.89  |
| Male (n=9261)                  | 7526  | 81.26 | 1735 | 18.734 |
| $\leq 2$ MW (n=4651)           | 2989  | 64.26 | 1662 | 35.73  |
| <2 to 5 MW< (n=7182)           | 5788  | 80.59 | 1394 | 19.41  |
| $\leq 5$ MW (n=2635)           | 2295  | 87.09 | 340  | 12.90  |
| Absent (n=4167)                | 3401  | 81.61 | 766  | 18.38  |

Source: The author, 2023.

The number of walking trips per zone is shown in Figure 26. A pattern of a higher absolute number of walking trips conducted in the city center is prominent. Also, on the city's

lower-income peripheries, especially the extreme north and south. On the other hand, the highest income neighborhoods located in the southwest of the city have less reported walking.

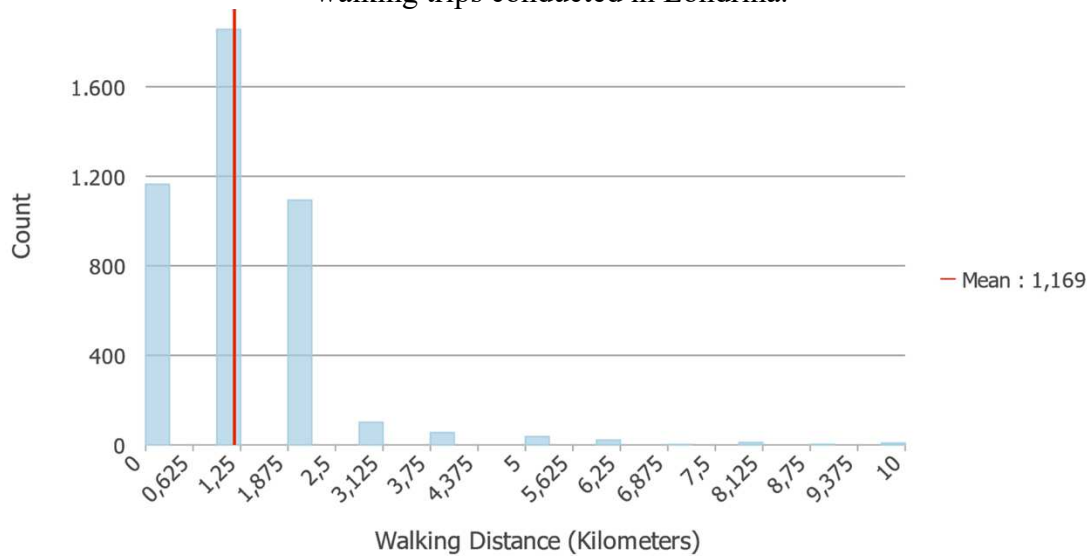
Figure 26 - Representation of the aggregate absolute number of walking trips conducted by zone.



Source: The author, 2023.

Figure 27 displays a histogram representing the walking distances of  $n=4162$  walking trips conducted in Londrina. The x-axis represents the walking distances in kilometers, while the y-axis indicates the frequency or number of walking trips falling within each distance bin. The histogram showcases a peak around 1.25 kilometers, indicating the highest concentration of walking trips within this distance range. This suggests that a significant number of individuals in the study opted for relatively short walking distances during their trips in Londrina. The average length of a walking trip, calculated to be 1.169 kilometers, serves as a central reference point for the distribution. This indicates that, on average, the walking trips were slightly shorter than 1.25 kilometers, as seen from the histogram's peak.

Figure 27 – Histogram of walking distances of the n=4162 walking trips conducted in Londrina.

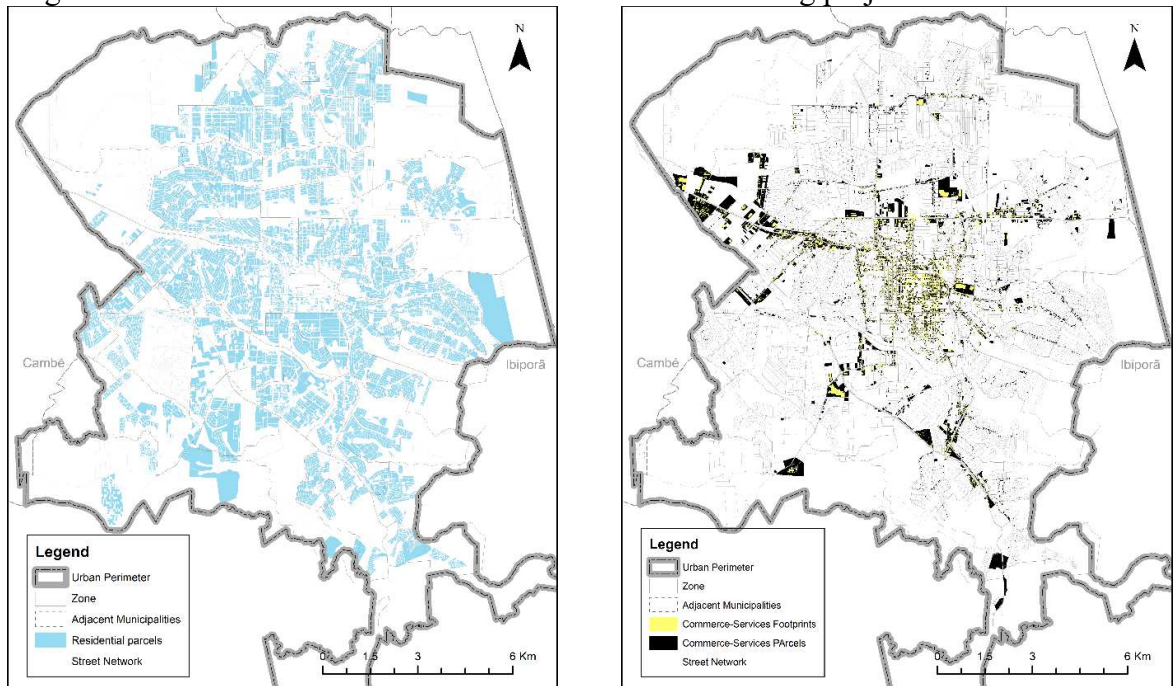


Source: The author, 2023. Based on data from Londrina's administration (2018).

#### 4.3.1 Macroscale Walkability for Londrina

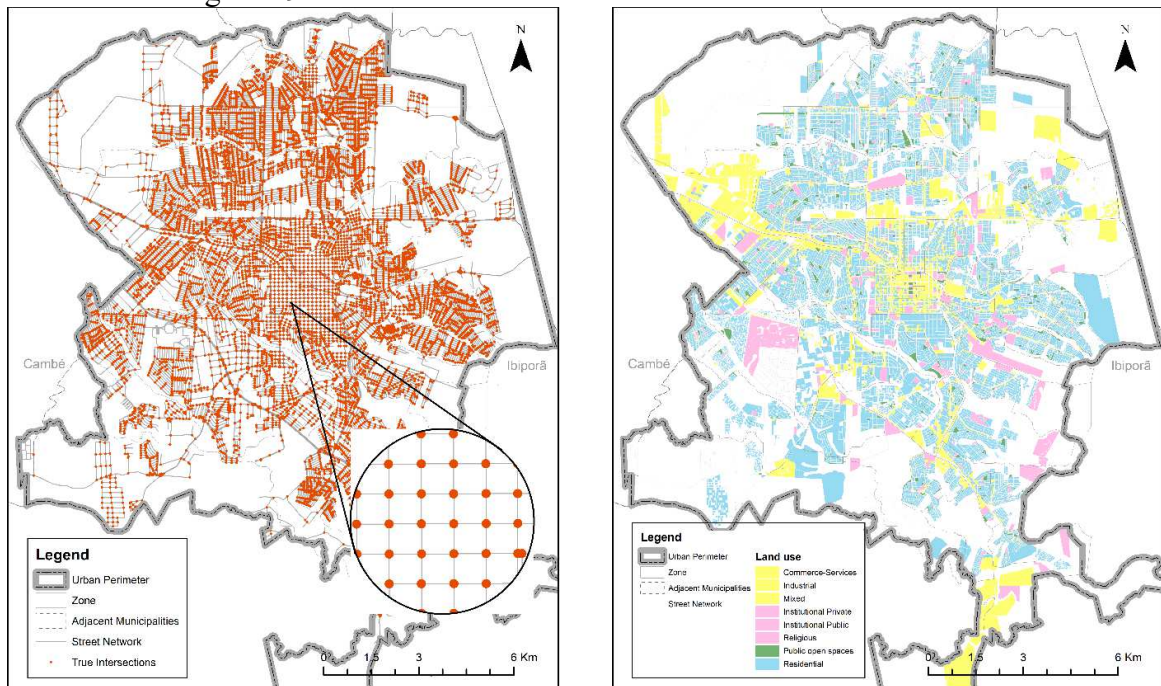
Figures 28 to 30 illustrate five separate variables employed in assessing the macro-level walkability in Londrina-PR. This data encompasses details about residential lots, commercial construction area, intersection, land use data, and street integration computations derived from space syntax. The residential use data delivers specifics on the geographical dispersion in residential zones in Londrina-PR. The commercial built-up area data concentrates on the physical range of edifices utilized for business activities in the city. It allows for the evaluation of spatial dispersion and concentration of commercial buildings as well as to the built area vs designated lot area to this use. The intersection data represents the positions and features of intersections within the road network of Londrina-PR. This information permits the analysis of street interconnectivity, traffic movement, and the efficacy of the transportation network. The land use data offers insights into the diverse categories of land usage in Londrina-PR, including residential, commercial, industrial, and recreational zones. It aids in comprehending the dispersion and makeup of different land uses across the city. Lastly, the street integration computations from space syntax employ analysis methods to appraise street integration metrics for the street network of Londrina-PR.

Figure 28 - Residential and commercial lots and their building projections from Londrina.



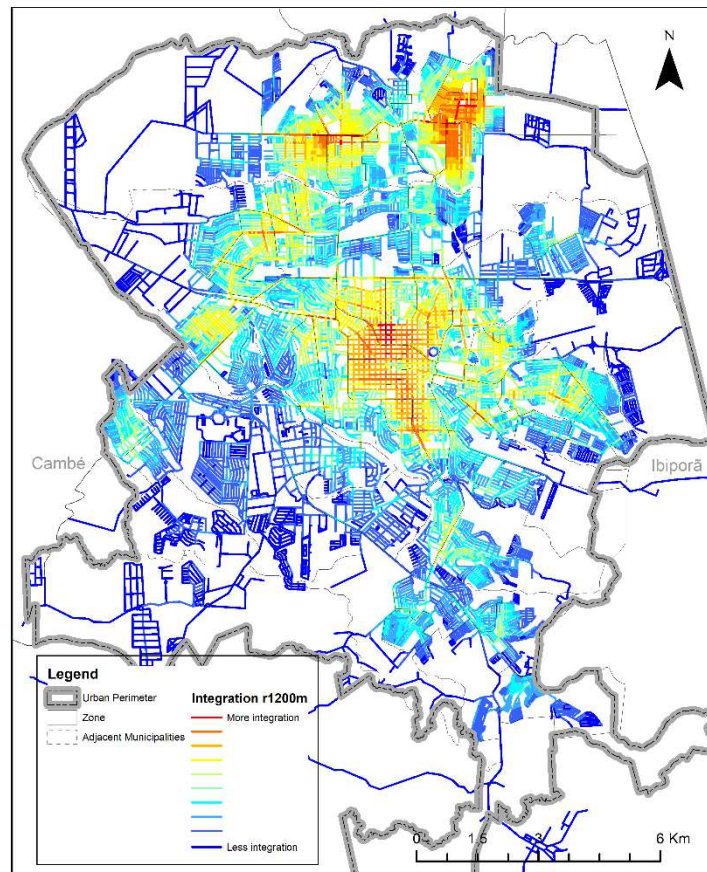
Source: IPPUL (2019), Elaborated by the author, 2021.

Figure 29 - True Intersections and land use data from Londrina.



Source: IPPUL (2019), Elaborated by the author, 2021.

Figure 30 - Street Integration at the 1200-meter radius.



Source: IPPUL (2019), Elaborated by the author, 2021.

The multilevel logistic regression results for the macroscale variable modeling are shown in Table 17. The ICC for the null model predicting the odds of walking 29.88%, indicating that a relatively high proportion of outcome variation can be attributed to differences between zones. The smallest AIC was considered for identifying the best-fitted model, as it intends to prevent the inclusion of irrelevant predictors. However, it must be emphasized that the number in itself is not meaningful as it is used to compare but is not interpretable on its own.

The initial one contains the built environment variables as explanatory of walking trips being made versus not. A second model contains all individual factors and all macro-built environment variables. Lastly, the best-fitted model is presented, considering all individual factors and built environment variables at the macro scale. Such variables were the only suited to compose the model, showing significance, and having their inclusion beneficial for the model quality (analyzed through AIC).

Table 17 - Multilevel logistic regression analysis for the relationship between the choice of walking, built environment, and individual factors (n = 18635).

| Model Predictor  | (1a) BE variables                     |                |          | (1b) IBGE variables                   |                |         | (2) Complete model                    |                |          | (3) Best fitted model                |                |          |
|--|---------------------------------------|----------------|----------|---------------------------------------|----------------|---------|---------------------------------------|----------------|----------|--------------------------------------|----------------|----------|
|  | OR                                    | 95% CI         | p-value  | OR                                    | 95% CI         | p-value | OR                                    | 95% CI         | p-value  | OR                                   | 95% CI         | p-value  |
| <b>Individual</b>  |                                       |                |          |                                       |                |         |                                       |                |          |                                      |                |          |
| <i>Age</i>   |                                       |                |          |                                       |                |         |                                       |                |          |                                      |                |          |
| Child (≤12)  | -                                     | -              | -        | -                                     | -              | -       | 1.00*                                 | -              | -        | 1.00*                                | -              | -        |
| Adult (<13 to 59<)   | -                                     | -              | -        | -                                     | -              | -       | 0.567                                 | 0.507<br>0.634 | <0.001** | 0.567                                | 0.507<br>0.634 | <0.001** |
| Elderly (≤60)  | -                                     | -              | -        | -                                     | -              | -       | 0.559                                 | 0.491<br>0.637 | <0.001** | 0.559                                | 0.491<br>0.636 | <0.001** |
| <i>Gender</i>  |                                       |                |          |                                       |                |         |                                       |                |          |                                      |                |          |
| Female - Reference   | -                                     | -              | -        | -                                     | -              | -       | 1.00*                                 | -              | -        | 1.00*                                | -              | -        |
| Male   | -                                     | -              | -        | -                                     | -              | -       | 0.741                                 | 0.687<br>0.800 | <0.001** | 0.742                                | 0.688<br>0.801 | <0.001** |
| <i>Household income</i>  |                                       |                |          |                                       |                |         |                                       |                |          |                                      |                |          |
| ≤ 2 MW   | -                                     | -              | -        | -                                     | -              | -       | 1.00*                                 | -              | -        | 1.00*                                | -              | -        |
| <2 to 5 MW<  | -                                     | -              | -        | -                                     | -              | -       | 0.474                                 | 0.432<br>0.520 | <0.001** | 0.47                                 | 0.432<br>0.520 | <0.001** |
| ≤ 5 MW   | -                                     | -              | -        | -                                     | -              | -       | 0.335                                 | 0.288<br>0.390 | <0.001** | 0.33                                 | 0.287<br>0.388 | <0.001** |
| Absent   | -                                     | -              | -        | -                                     | -              | -       | 0.467                                 | 0.418<br>0.521 | <0.001** | 0.46                                 | 0.418<br>0.520 | <0.001** |
| <i>Employment</i>  |                                       |                |          |                                       |                |         |                                       |                |          |                                      |                |          |
| Formal   | -                                     | -              | -        | -                                     | -              | -       | 1.00*                                 | -              | -        | 1.00*                                | -              | -        |
| Informal   | -                                     | -              | -        | -                                     | -              | -       | 1.700                                 | 1.501<br>1.926 | <0.001** | 1.701                                | 1.502<br>1.926 | <0.001** |
| Unemployed and Refusal   | -                                     | -              | -        | -                                     | -              | -       | 3.259                                 | 2.954<br>3.595 | <0.001** | 3.259                                | 2.955<br>3.596 | <0.001** |
| <b>Built environment (cont.var/ average z-scores per zone)</b> |                                       |                |          |                                       |                |         |                                       |                |          |                                      |                |          |
| Land use Mix (entropy)   | 1.062                                 | 0.878<br>1.284 | 0.538    | 1.004                                 | 0.942<br>1.070 | 0.902   | 1.184                                 | 0.986<br>1.421 | 0.069    | 1.257                                | 1.085<br>1.456 | <0.001** |
| Residential Density  | 1.069                                 | 0.876<br>1.305 | 0.509    | 1.000                                 | 1.000<br>1.000 | 0.617   | 1.030                                 | 0.854<br>1.244 | 0.750    | -                                    | -              | -        |
| Intersection Density   | 1.184                                 | 0.934<br>1.500 | 0.162    | 1.277                                 | 1.086<br>1.501 | 0.003   | 1.095                                 | 0.876<br>1.369 | 0.424    | -                                    | -              | -        |
| Retail FAR   | 1.049                                 | 0.864<br>1.275 | 0.626    | 1.063                                 | 0.906<br>1.248 | 0.453   | 1.068                                 | 0.888<br>1.284 | 0.480    | -                                    | -              | -        |
| Integration r1200  | 1.004                                 | 1.000<br>1.008 | <0.001** | 1.004                                 | 1.000<br>1.008 | 0.057   | 1.005                                 | 1.001<br>1.008 | 0.005**  | 1.006                                | 1.003<br>1.009 | <0.001** |
| <b>Model Fit</b>   | <b>AIC= 19122.5, logLik = -9554.2</b> |                |          | <b>AIC= 19123.1, logLik = -9554.6</b> |                |         | <b>AIC =17394.2, logLik= -8682.1-</b> |                |          | <b>AIC =17390.4, logLik= -8683.2</b> |                |          |

Note<sup>1</sup>: \*=1,00: reference category; CI 95%: 95% Confidence Interval. \*\* significative at the α=0,05 level. \*\*\* ANOVA's Pr(>Chisq) at α=0.05

Source: The author, 2023.

The relationship between the choice of walking versus using other modes of transportation and the macroscale characteristics of the built environment, moderated by individual factors, was examined through multilevel logistic regression modeling. The first model was characterized by the quantification of the predictive capacity of the built environment variables, which was achieved using data obtained from the Urban Planning

Institute of Londrina to measure mixed land use in relation to walking. Notably, integration within a 1200-meter radius emerged as the sole significant variable, with the OR proving challenging to interpret. The second model involved only macro-scale-built environment variables, relying on the BE variables as measured by the data from IBGE's National Registry of Addresses for statistical purposes. In this context, no BE variable exhibited significant contribution to the prediction of walking.

In the third model, due to the observed limited efficacy of replacing the data for measuring mixed land use from locally collected data with those from IBGE, a comprehensive model was formulated utilizing land use data provided by IPPUL. Relevance and significant contribution to the prediction of walking were established for all individual variables. However, among the built environment variables, only integration within a 1200-meter radius was identified as relevant but with uninterpretable OR.

Finally, the fourth model took into account individual variables based on the information derived from the OD survey conducted for the city of Londrina. Factors such as self-reported age, gender, family income, and employment status were considered. In the best-fitted model, all individual factors proved relevant. When it comes to individual factors, all variables were found to be significant. From the categories in Age, the elderly ( $\leq 60$ ) and adults ( $< 13$  to  $59 <$ ) showed less probability of a trip being made by foot. Compared to the reference class of children under 12 years old, adults and elderly individuals demonstrated lower odds of undertaking a walking trip versus using alternative means of transport, with ORs of 0.56 and 0.55, respectively. Where for each trip made by and child ( $\leq 12$ ), 0.56 and 0.55 trips were made by adults and the elderly, respectively.

Regarding gender, the OR for an individual reporting as male to walk, compared to the reference class of females, was 0.74. This suggests that men or individuals reporting as male have lower odds of walking compared to those reporting as female. When it comes to gender, women have presented higher OR of walking than men, for each trip made by a woman, 0.74 trips were made by men.

With respect to income, using as the reference class individuals reporting family income of less than or equal to two MW, it was identified that all other classes, two to five MW, more than five MW, or those who refused to answer, had lower odds of walking than individuals of lower income or from the lowest socioeconomic stratum, with ORs of 0.47, 0.33, and 0.46, respectively. Thus, they had less than half the odds of walking compared to individuals of lower income classes. In relation to employment status, with the reference class being those with a

formal job, individuals with informal jobs or those reporting unemployment, or who refused to comment on their employment status, had substantially higher odds of walking compared to those with formal jobs, with respective ORs of 1.7 and 3.25.

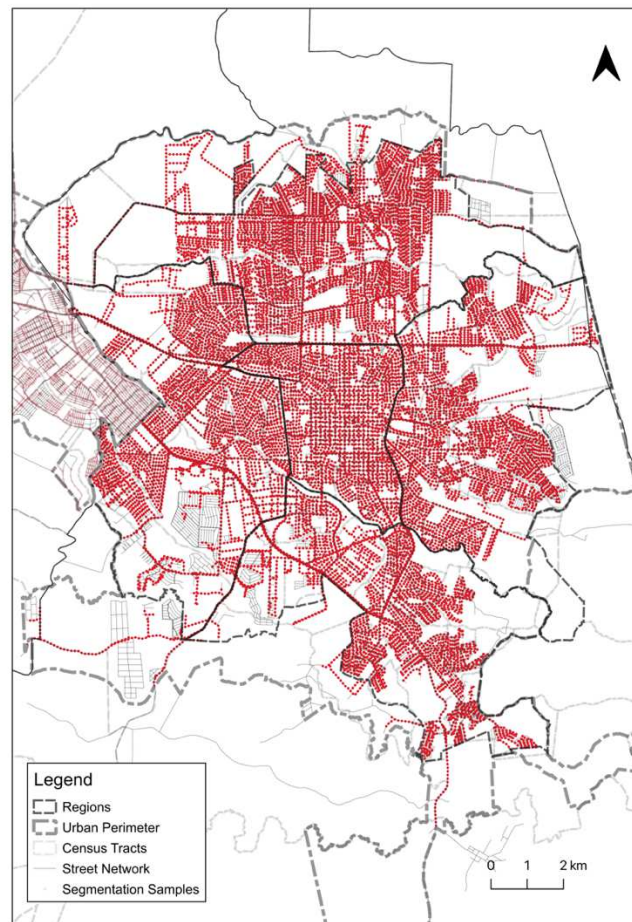
Within the best-fitted model concerning built environment variables at a larger scale, it was identified that land use mix was significant for the model moderated by individual variables. Individuals residing in zones with higher values of mixed land use have higher OR of undertaking a walking trip versus using other modes of transportation, with an OR of 1.25. Previous literature, including work conducted in Brazil (Leão, 2019), indicates entropy/land use mix as the most critical variable in the influence of walking behaviors. Furthermore, the variable of integration proved relevant for the model; however, its ORs, being very close to 1, are not interpretable. Thus, it cannot be concluded that higher or lower levels of integration, in this case, and with this sample, are relevant for promoting walking in the presence of these individual variables and other environmental variables at the macro scale.

#### **4.3.2 Microscale Walkability for Londrina**

From the sampling conducted considering points 100 meters apart through the street network in the ArcGIS 10.6 software, n=19320 initial the geo-tagged sample points were selected. In the acquisition procedure from GSV, a total of n=17860 panoramic images were successfully retrieved from the API totaling n=71440, as each panorama is composed of 4 images. It is relevant to point out that the sampling is proportional to the size and extent of each zone's street network (Figure 31).

After the processing of the pretrained model, the inferences on the categories of pixels that composed the images from the selected sample points were the bases for the following analysis. The absolute count of pixels, a ratio, and a z-score of each segment class for each image was obtained. Without aggregation, the results indicate the most predominant characteristics that compose the streetscapes from the entire city.

Figure 31 - Sampling of geo-tagged points for image extraction.



Source: The author, 2023.

An initial statistical analysis was conducted with the average values of z-scores from semantic segmentation results aggregated in zones and the absolute number of walking trips per zone. Spearman's rho correlation coefficients were obtained and are depicted in Figure 32. Hypothesis testing was conducted; therefore, coefficients without significance at the  $\alpha=0,05$  level are not displayed. Despite representing linear relationships between variables and the sample size limited to the number of zones ( $n=83$ ), it is possible to interpret results tentatively.

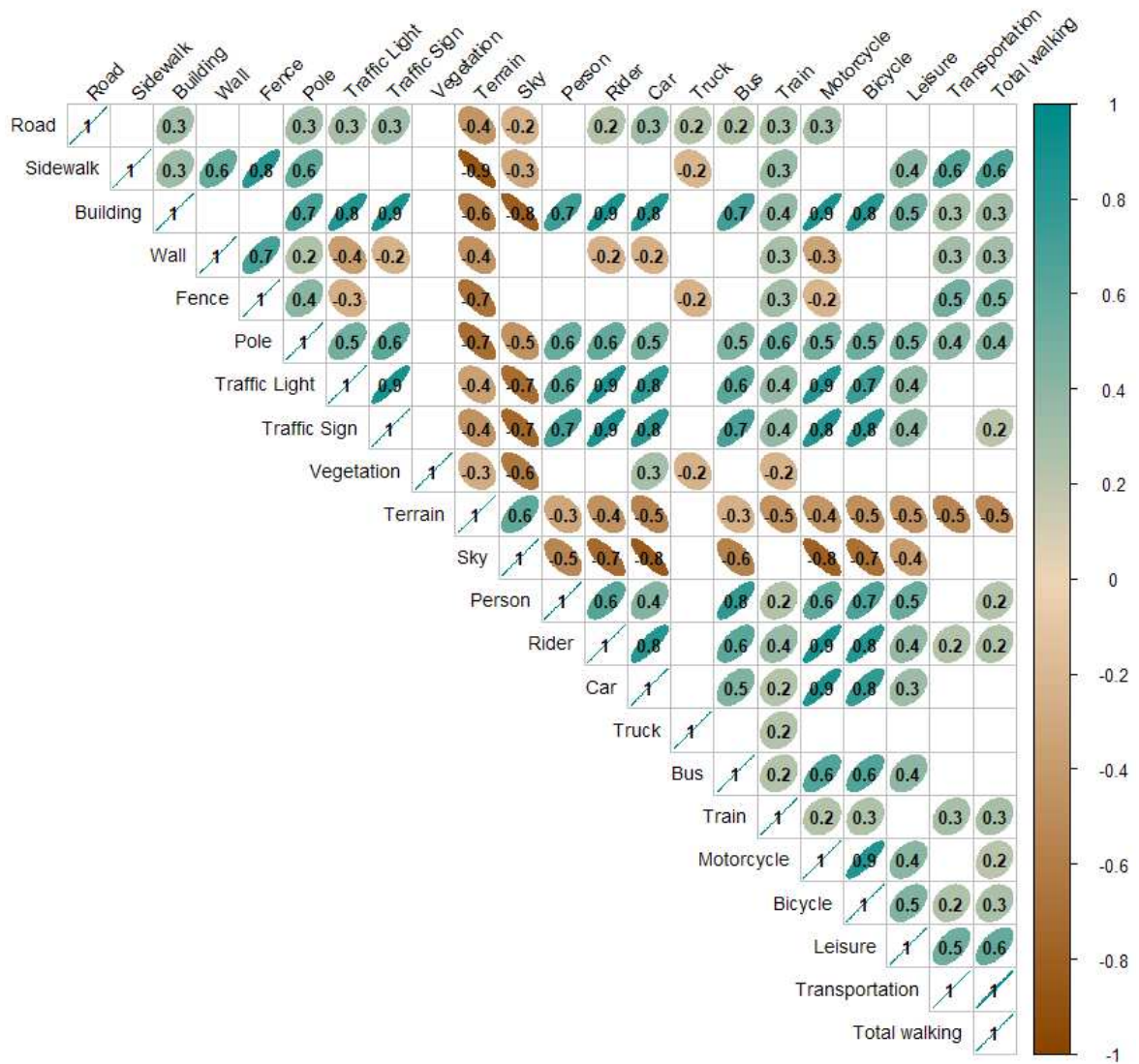
Through a parametric correlation statistical analysis, several associations between the built environment and walking behaviors were identified. The analysis considered microscale variables for total walking and differentiated between walking for leisure and transport. It was found that leisure walking is inversely related to the presence of vacant lots and higher sky visibility. This is consistent with literature suggesting vacant lots may elicit perceptions of insecurity, discouraging pedestrian activity. In addition, high sky visibility could suggest a lack of elements, such as trees or buildings, that provide a comforting sense of enclosure, a feature valued by pedestrians and frequently discussed in classic urban design theory.

The Terrain category relates to grass, horizontal vegetation, soil, or sand, represented mainly in the case study by empty plots, quite present in some city areas like its extreme outskirts and vertical residential neighborhoods under development. Sky relates to the existence of an open sky. This category presented a slightly stronger negative relationship to walking. This might relate to the articulation of numerous urban design theorists (Alexander et al., 1977; Cullen, 1961; A. B. Jacobs, 1993) on the sense of Enclosure. Simply put, it indicates that people behave favorably to fixed boundaries as being safe and providing a sense of position, making outdoor spaces seem room-like (R. Ewing & Handy, 2009). In general, large amounts of the visible sky from a streetscape perspective exist in this case study, which might inhibit walking.

Contrastingly, leisure walking is significantly correlated with the presence of buildings, signposts, people, and bicycles. These features contribute to an environment's interest and perceived safety, enhancing the vibrancy of the area and making it more conducive to leisure walking. The 'eyes on the street' effect due to the presence of people and bicycles provides a safer environment for such activity. The correlations found between Buildings with the absolute number of walking trips per zone can be considered moderately strong positive values. Despite not accounting for the different types of land use mixtures, a widely recognized positive impact in walking (Christian et al., 2011; Rachele et al., 2018; Sugiyama et al., 2019), the sole existence of buildings (Building, skyscraper, house, bus stop building, garage, carport) was, in this specific case, a strong linear predictor of walking.

When it comes to transport walking, even stronger correlations were identified with the presence of sidewalks and street trees. Sidewalks provide a designated safe space for pedestrians, separated from vehicular traffic, promoting more utilitarian walking behavior. Street trees contribute to environmental aesthetics and provide shade, showing a moderate correlation with transport walking. Finally, the results suggest that overall walking is positively associated with the presence of sidewalks and fences, which provide defined and secure paths, thereby encouraging walking. However, it's less likely in areas with vacant lots, possibly due to perceived insecurity and lack of surveillance in such areas.

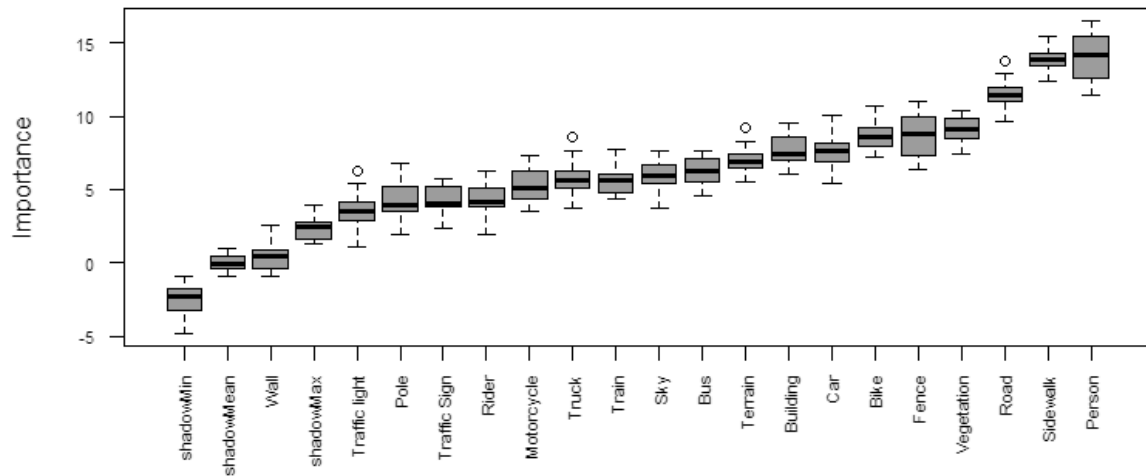
Figure 32 - Spearman’s rho correlation between absolute walking trips per zone and identified category values per zone in Londrina.



Source: The author, 2023.

The second analysis between micro-scale characteristics and walking was conducted, controlling for sociodemographic characteristics. A feature selection procedure was conducted due to the large number of categories. Walking data was used not aggregated; therefore, the sample is related to the total number of trips (n = 18635) in binary form and their respective zone’s micro-scale characteristics. The method Boruta was applied in All categories obtained in semantic segmentation were considered for the feature selection. A boxplot depicting their relative importance is shown in Figure 33. This order was considered for the following regression modeling.

Figure 33 - Boxplot representing the importance of micro-scaled categories obtained from semantic segmentation to walking (or not). Londrina



Source: The author, 2023.

The results from the second multilevel logistic regression analysis are presented in Table 18. ICC for the null model, which predicts the odds of walking versus not walking, is 29.88%. This statistic represents the proportion of outcome variance attributable to differences between zones. The maximum VIF appears in the model considered best fitted. Despite the high VIF, it remains below 8, suggesting an acceptable level of multicollinearity among the explanatory variables. Thus, it is understood that the interpretability of the results can be maintained.

In this case, the AIC cannot be the sole comparison metric for identifying the best-fitted model, given that the VIF of the complete model is extremely high, despite being lower than that of the well-adjusted model. In this context, the model considered as best fitted does not present a superior AIC, but it was achieved through the inclusion and exclusion of variables until the lowest value of variable inflation was found, given that none of the combinations could maintain a better AIC as presented by the complete model.

In adjusting the multilevel logistic regression models, microscale variables and individual factors were considered as predictors for the choice of walking in Londrina. A preliminary model composed solely of built environment variables identified several significant variables. However, this model showed a high AIC and a high inflation value. A second, "complete" model included both individual variables (age, reported gender, family income, employment status) and microscale environmental variables. In this model, all individual factors were relevant, and most environmental variables were significant. Nonetheless, the inflation value for the complete model increased exponentially compared to the previous model, indicating a high level of collinearity among the variables.

Ultimately, the best-fitted model suggested that all individual variables were relevant. When considering the class of children aged 12 or less as the reference class, adults and seniors had lower odds of making a trip, with ORs of 0.59 for both categories. Thus, for every trip made by a child, 0.59 trips were made by adults or seniors. Regarding reported gender, with individuals who reported being female as the reference class, those who reported being male had odds of 0.73 for making a trip on foot. Thus, for every trip made by an individual who reported being female, 0.73 trips were made by one who reported being male.

In terms of family income, similar to previous analyses, using the class of two MW or less as the reference class, it was identified that the higher the family income, the lower the odds of an individual making a trip on foot. For every trip made on foot by an individual from the socioeconomic class with a family income of two MW or less, 0.48 trips were made by individuals from the class of two to five MW, and 0.38 trips were made by individuals from the class of five MW or more. Finally, for individual characteristics, the employment status compared to individuals with formal jobs indicated that those with informal jobs had odds of 1.64, and those unemployed or who refused to answer had odds of 1.6, indicating a significant increase in the odds of an individual walking if they have an informal job or are unemployed.

In the best-fitted model concerning environmental variables, the presence of people, roads, and vegetation was associated with reduced odds of an individual choosing to walk. In contrast, the presence of sidewalks increased the likelihood, or the odds, of walking. When examining the residential zone of an individual, those living in areas with fewer trees, fewer paved streets, and fewer people tend to walk less. However, areas with a more significant sidewalk presence show higher odds of walking, with an odds ratio (OR) of 5.23. The findings related to vegetation warrant discussion. While this model suggests vegetation reduces the likelihood of walking, when considered alongside other variables, numerous studies, including Dadvand et al. (2016), have consistently highlighted vegetation's positive impact on walking, emphasizing the association between urban greenness and physical activity.

Table 18 - Multilevel logistic regression analysis for the relationship between the choice of walking, micro scale-built environment, and individual factors (n = 18635).

| Model Predictor  | OR                                     | 95% CI               | p-value      | OR                                     | 95% CI          | p-value      | OR                                 | 95% CI         | p-value  |
|--|--|----------------------|--------------|--|-----------------|--------------|------------------------------------|----------------|----------|
|  | (1) BE Model                           |                      |              | (2) Complete model                     |                 |              | (3) Best fitted model              |                |          |
| <b>Individual</b>  |  |                      |              |  |                 |              |                                    |                |          |
| <b>Age</b>   |  |                      |              |  |                 |              |                                    |                |          |
| Child(≤12)   | -                                      | -                    | -            | 1.00*                                  | -               | -            | 1.00*                              | -              | -        |
| Adult (<13 to 59<)   | -                                      | -                    | -            | 0.607                                  | 0.542<br>0.681  | <0.001**     | 0.598                              | 0.534<br>0.670 | <0.001** |
| Elderly (≤60)  | -                                      | -                    | -            | 0.612                                  | 0.536<br>0.699  | <0.001**     | 0.598                              | 0.524<br>0.682 | <0.001** |
| <b>Gender</b>  |  |                      |              |  |                 |              |                                    |                |          |
| Female   | -                                      | -                    | -            | 1.00*                                  | -               | -            | 1.00*                              | -              | -        |
| Male   | -                                      | -                    | -            | 0.732                                  | 0.677<br>0.790  | <0.001**     | 0.733                              | 0.678<br>0.791 | <0.001** |
| <b>Household income</b>  |  |                      |              |  |                 |              |                                    |                |          |
| ≤ 2 MW -   | -                                      | -                    | -            | 1.00*                                  | -               | -            | 1.00*                              | -              | -        |
| <2 to 5 MW<  | -                                      | -                    | -            | 0.485                                  | 0.442<br>0.533  | <0.001**     | 0.481                              | 0.438<br>0.528 | <0.001** |
| ≤ 5 MW   | -                                      | -                    | -            | 0.354                                  | 0.304<br>0.414  | <0.001**     | 0.348                              | 0.298<br>0.405 | <0.001** |
| Absent   | -                                      | -                    | -            | 0.474                                  | 0.424<br>0.529  | <0.001**     | 0.472                              | 0.422<br>0.526 | <0.001** |
| <b>Employment</b>  |  |                      |              |  |                 |              |                                    |                |          |
| Formal   | -                                      | -                    | -            | 1.00*                                  | -               | <0.001**     | 1.00*                              | -              | -        |
| Informal   | -                                      | -                    | -            | 1.634                                  | 1.440<br>1.853  | <0.001**     | 1.640                              | 1.446<br>1.859 | <0.001** |
| Unemploy ed and Refusal  | -                                      | -                    | -            | 3.161                                  | 2.861<br>3.492  | <0.001**     | 3.162                              | 2.864<br>3.491 | <0.001** |
| <b>Built environment (cont.var/ average z-scores per zone)</b> |  |                      |              |  |                 |              |                                    |                |          |
| Bike   | 0.117                                  | 0.285<br>47.667      | 0.000<br>*** | 9.826                                  | 2.218<br>43.534 | 0.002<br>**  | -                                  | -              | -        |
| Building   | 0.594                                  | 0.363<br>97.441      | 0.114        | 0.000                                  | 0.000<br>1.380  | 0.053        | -                                  | -              | -        |
| Bus  | 0.345                                  | 0.219<br>5.437       | 0.449        | 0.111                                  | 0.006<br>2.031  | 0.138        | -                                  | -              | -        |
| Car  | 0.316                                  | 0.308<br>3.250       | 0.078        | 0.000                                  | 0.000<br>0.368  | 0.034<br>*   | -                                  | -              | -        |
| Fence  | 0.626                                  | 0.961<br>4.073       | 0.086        | 0.000                                  | 0.000<br>0.508  | 0.038<br>*   | -                                  | -              | -        |
| Motocycle  | 0.259                                  | 0.376<br>1.780       | 0.169        | 0.293                                  | 0.038<br>2.240  | 0.236        | -                                  | -              | -        |
| Person   | 0.402                                  | 0.147<br>1.102       | 0.076        | 0.301                                  | 0.103<br>0.877  | 0.027<br>*   | 0.854                              | 0.786<br>0.927 | <0.001** |
| Pole   | 0.215                                  | 0.984<br>0.469       | 0.014<br>*   | 0.021                                  | 0.000<br>0.575  | 0.021<br>*   | -                                  | -              | -        |
| Rider  | 0.534                                  | 0.886<br>0.321       | 0.001<br>**  | 0.041                                  | 0.006<br>0.275  | 0.001<br>*** | -                                  | -              | -        |
| Road   | 0.132                                  | 0.121<br>14.362      | 0.072        | 0.000                                  | 0.000<br>0.053  | 0.033<br>*   | 0.355                              | 0.248<br>0.509 | <0.001** |
| Sidewalk   | 0.650                                  | 0.187<br>22.658      | 0.107        | 0.000                                  | 0.000<br>0.931  | 0.049<br>*   | 5.235                              | 4.170<br>6.572 | <0.001** |
| Sky  | 0.165                                  | 0.345<br>7.8761      | 0.105        | 0.000                                  | 0.000<br>0.604  | 0.047<br>*   | -                                  | -              | -        |
| Terrain  | 0.619                                  | 0.140<br>27.425      | 0.079        | 0.000                                  | 0.000<br>0.098  | 0.036<br>*   | -                                  | -              | -        |
| `Traffic Light`  | 0.558                                  | 0.172<br>18.103      | 0.004<br>**  | 3.810                                  | 1.106<br>13.118 | 0.033<br>*   | -                                  | -              | -        |
| `Traffic Sign`   | 0.807                                  | 0.206<br>0.031       | 4.41e<br>*** | 0.008                                  | 0.001<br>0.035  | 0.000<br>*** | -                                  | -              | -        |
| Truck  | 0.235                                  | 0.536<br>10.271      | 0.226        | 0.008                                  | 0.000<br>5.154  | -            | -                                  | -              | -        |
| Vegetation   | 0.128                                  | 0.261<br>6.3131<br>2 | 0.092        | 0.000                                  | 0.000<br>0.157  | 0.1437       | 0.349                              | 0.271<br>0.449 | <0.001** |
| Wall   | 0.222                                  | 0.961<br>5.123       | 0.089        | 0.000                                  | 0.000<br>0.477  | 0.0418<br>*  | -                                  | -              | -        |
| <b>Model Fit</b>   | <b>AIC=18578.4,<br/>logLik=-9269.2</b> |                      |              | <b>AIC=17019.2,<br/>logLik=-8482.6</b> |                 |              | <b>AIC=17135,<br/>logLik=-8554</b> |                |          |

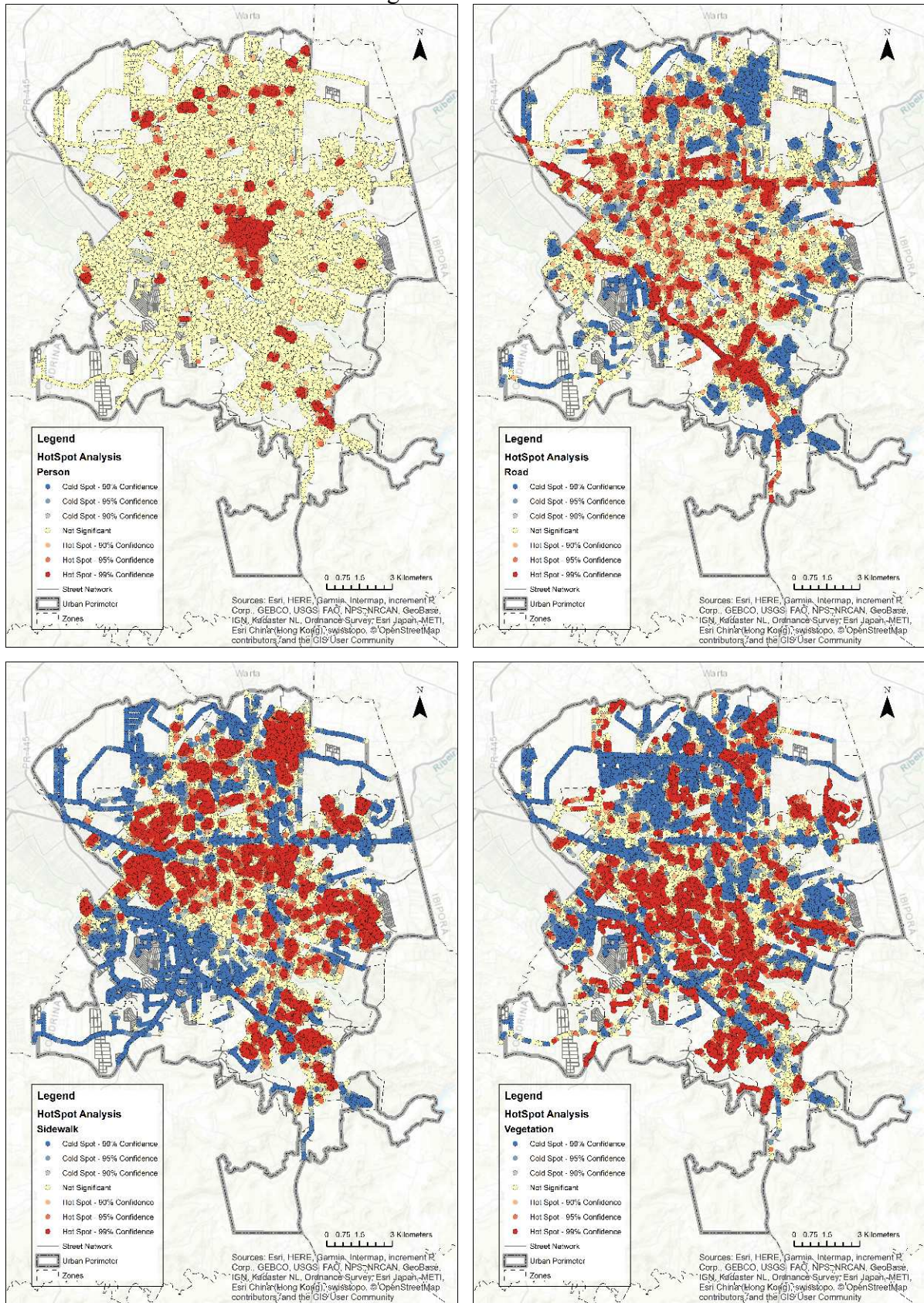
Note<sup>1</sup>: \*=1,00: reference category; CI 95%: 95% Confidence Interval. Note<sup>2</sup>: \*\* significant at the  $\alpha=0,05$  level; \*\*\* significant at the  $\alpha=0,01$  level

Source: The authors, 2023

These ideas can be further analyzed in Figure 34, which depict Hotspots analyses for the distribution of the most relevant semantic segmentation categories in the regression model presented in Table 18. The 'Person' category reveals hotspots primarily in the commercial or mixed-use areas of the city, consistent with traditional urban planning theories (J. Jacobs, 1961). However, OR for People in the model is 0.85. This suggests that individuals living in more densely populated areas have a lower likelihood of walking. Given the hotspot concentration and the conflicting model results, it might be beneficial to introduce macro-scale variables into the model to verify this discrepancy.

The insights regarding Sidewalks become more comprehensible when examined alongside spatial analysis. Take, for instance, the southwestern sector of the city, which is notably affluent. In this area, there is a notable concentration of sidewalk Coldspots juxtaposed with road hotspots. Drawing from the individual variable results, there is a trend: as income increases, the propensity to walk decreases. The Roads category, as per the training database, encompasses not only conventional car lanes but also car parking spaces, and roundabouts. When one considers that lower-income individuals tend to walk more frequently, it becomes evident that an increase in road infrastructure correlates with decreased pedestrian activity. More roads often lead to enhanced vehicular traffic, making walking less appealing or, in some cases, less safe. As for the Vegetation category, the higher the vegetation score (which accounts for features like trees and vertical green fences), the less likely an individual is to walk. This could suggest that while vegetation is aesthetically pleasing, dense vegetation might create barriers or perceived safety concerns for pedestrians, discouraging walking. Such a phenomenon warrants further discussion conducted in the next chapter.

Figure 34 - Hot Spot for the distribution of the Person, Road, Sidewalk and Vegetation categories in Londrina.



Source: The author, 2023.

#### 4.4 Results and discussions of multiscale modelling in Rolândia, Cambé and Londrina

Table 19 depicts the three different models identified as the best fit for each selected case study city. These models were created using available secondary data for each city; therefore, they do not present the same individual variables. Nonetheless, a certain level of comparability among the models and cities is maintained. Modeling, incorporated variables from both micro and macro scales. These include the street environment elements as well as larger scale ones, related to Urban Design and Urban Planning fields.

In the case of Londrina, the best fit model demonstrated a VIF of 10, qualifying it as the optimal configuration. Despite the decrease in the VIF to 8 following the removal of entropy, the interpretability of the model was deemed to remain intact. For the Rolândia model, the VIF of the best fit was observed to be less than 4. Concurrently, the VIF values for the Cambé model were found to be under 3.

In all models presented in Table 19, despite the differences among them, all individual variables were important for the modeling and exhibited significance. Concerning age, both the cities of Rolândia and Londrina provided sufficient data to consider this variable. In Rolândia, with children aged 12 or younger as the reference group, it was identified that adults have lower odds of walking (OR=0.53), whereas older adults have higher odds (OR=1.73) of walking compared to children and adults. In contrast, in Londrina, with the same reference group, both adults and older adults showed lower odds of walking versus other modes of transportation, with ORs of approximately 0.59 for both groups. This implies that for every walking trip undertaken by a child, around half a trip would be undertaken by an adult or older adult in the city of Londrina.

Adults, in general, have higher access and options for transportation, including individual motorized transport than the elderly over 60 and children under 12, activities that do not necessarily demand motorized transportation might be performed by foot. Therefore, the results obtained do not necessarily agree with high income country literature evidence that indicates a decrease in walking levels as age increases (Van Cauwenberg et al., 2012). Considering these results comparatively, it can be interpreted that in a smaller city like Rolândia, there might be a greater capacity for 'aging in place' in a more active manner, which requires higher levels of social capital and social networks (Pani-Harreman et al., 2021) that can be facilitated through shorter distances, compactness. Further, the under-sampling (Table 15) in the elderly and child categories can have played a role in this result.

In relation to gender, only the cities of Rolândia and Londrina provided this information for modeling from secondary data. For both cities, with individuals who reported being female as the reference group, individuals who reported being male have lower odds of walking. For Rolândia, there was a significant group who refused to provide information or had missing data; this group showed significance in the prediction of the best fit model. Compared to the female gender, this group displayed higher odds of walking. Conjectures that women may walk more due to disparities in income and their imposed social roles that account more for domestic burdens such as taking children to school are possible. This aligns with the general understanding that women have more complex travel patterns and due to social and economic restraints are still bound to walk more for domestic and childcare, even as they participate in the workforce (Schwanen et al., 2008). However, walking for leisure is significantly below recommendations for women (Adlakha & Parra, 2020), which might relate to the fear promoted by social conditions and public space design that imposes notions on women's roles and appropriate places (R. Pain, 2001). In LMICs, a significant gender gap exists in physical activity, with women's inactivity rates surpassing men's at 31.7% despite the known health benefits of PA (Adlakha & Parra, 2020). Despite persistent absence of evidence linking these patterns in countries like Brazil (Leão & Kanashiro, 2021). It can be conjectured that, in this scenario, walking for leisure is a generally suppressed behavior among women, being avoided when income, education, and employment status condition for other opportunities.

In terms of family income, the information used for the cities of Londrina and Rolândia was based on self-reported income in the origin-destination questionnaire applied in both cities. For the city of Cambé, this information was not part of the applied questionnaire in a manner sufficient for systematic use. Therefore, the nominal monthly income per building or residence, attributed by the IBGE in the 2010 census (IBGE, 2010) to census tracts of the city, was used to represent family income.

For the city of Rolândia, with the group earning 2 MW or less as a reference, individuals earning between 2 to 5 MW (OR=0.35) and more than 5 MW were found to have lower odds of walking. It is important to note that the group earning more than 5 MW likely did not show significance in the model, possibly due to the small sample size. On the other hand, individuals who did not answer or had missing information regarding self-reported family income had substantially higher odds of walking, even more than individuals in the lowest income class. In Cambé, also with the lowest income class as a reference, individuals earning between 2 to 5 MW (OR=0.65) had lower odds of walking than those earning less than 2 MW. Lastly, in the

city of Londrina, with individuals receiving a family income of less than 2 MW as reference, those earning between 2 to 5, more than 5 MW, and those with missing information or who refused to answer, had lower odds of walking than those in the lowest socioeconomic level, with ORs of 0.47, 0.34, and 0.46 respectively. Generally, this result is in line with the literature that indicates that individuals from socioeconomically disadvantaged groups walk more for transportation (Yang et al., 2012). However, they are also disproportionately affected by adverse health outcomes (Hilland et al., 2020).

The household income and employment status results are largely in consonance with the literature indicating that the lowest income categories have higher chances of walking, especially for transportation. The demonstrated relationship between employment status and walking, can be due to those with informal jobs or unemployed are more inclined to walk due to economic considerations. As walking is a cost-free mode of transportation, it can be an attractive option for individuals looking to mitigate expenses associated with other modes of transport, such as cars or public transit. Geographical considerations can also factor into this propensity for walking.

Consideration must be given to the fact that individuals with lower income, and possibly those reporting informal employment status or unemployment, might reside in areas with higher concentrations of lower socioeconomic status individuals, and potentially less favorable environments for active travel. Additionally, there could be exposure to higher risk environments associated with informal or temporary jobs, which can tend to be physically intensive. These jobs can tend to be labor intensive, which often is linked to health burdens (Salvo et al., 2023), led by heightened exposure to pollution, crime risk, traffic, and even increases non-communicable diseases.

Finally, Londrina was the only city where information related to individuals' employment status was made available by the secondary data sources. With individuals with formal jobs as a reference, those with informal jobs and who are unemployed or refused to answer had substantially higher odds of walking, with OR of 1.64 and 3.16 respectively. In general, it is understood that individuals of lower income in Latin America, with informal employment or unemployed, are the ones who walk the most and for transportation, (Delclòs-Alió et al., 2022).

The same interpretation can be applied to gender, where women, or those who reported as female, consistently have higher odds of walking than their male counterparts. This might be attributable to the gap in income that allows men to have the financial power to

own a motor vehicle and to the gender disparities in women workforce participation still persistent in a country in development such as Brazil (Leão & Urbano, 2020). Regarding age, a difference was identified between the smallest and largest cities selected for this study, wherein the smallest city, older individuals displayed higher odds of walking than in the larger city.

Table 19 - Multilevel logistic regression analysis for the relationship between the choice of walking, micro and macro/macro built environment, and individual factors in Londrina., Cambé and Rolândia.

| Predictor  | (1) Best fitted Rolândia              |                 |                | (2) Best fitted Cambé                  |                |            | (3) Best fitted Londrina            |                |                |
|--|---------------------------------------|-----------------|----------------|--|----------------|------------|-------------------------------------|----------------|----------------|
|  | OR                                    | 95% CI          | p-value        | OR                                     | 95% CI         | p-value    | OR                                  | 95% CI         | p-value        |
| <b>Individual</b>                                    |                                       |                 |                |  |                |            |                                     |                |                |
| <b>Age</b>   |                                       |                 |                |  |                |            |                                     |                |                |
| Child ( $\leq 12$ )                                  | 1.00*                                 | -               | -              | 1.00*                                  | -              | -          | 1.00*                               | -              | -              |
| Adult ( $< 13$ to $59 <$ )                           | 0.532                                 | 0.365<br>0.776  | $< 0.001^{**}$ | -                                      | -              | -          | 0.598                               | 0.534<br>0.670 | $< 0.001^{**}$ |
| Elderly ( $\leq 60$ )                                | 1.733                                 | 1.060<br>2.836  | $< 0.05^*$     | -                                      | -              | -          | 0.594                               | 0.520<br>0.678 | $< 0.001^{**}$ |
| <b>Gender</b>  |                                       |                 |                |  |                |            |                                     |                |                |
| Female   | 1.00*                                 | -               | -              | 1.00*                                  | -              | -          | 1.00*                               | -              | -              |
| Male   | 0.532                                 | 0.365<br>0.776  | $< 0.001^{**}$ | -                                      | -              | -          | 0.73                                | 0.679<br>0.792 | $< 0.001^{**}$ |
| Absent   | 1.733                                 | 1.060<br>2.836  | $< 0.05^*$     | -                                      | -              | -          | -                                   | -              | -              |
| <b>Household income</b>                              |                                       |                 |                |  |                |            |                                     |                |                |
| $\leq 2$ MW -  | 1.00*                                 | -               | -              | 1.00*                                  | -              | -          | 1.00*                               | -              | -              |
| $< 2$ to $5$ MW $<$                                  | 0.354                                 | 0.138<br>0.908  | $< 0.05^*$     | 0.668                                  | 0.454<br>0.981 | $< 0.05^*$ | 0.477                               | 0.435<br>0.524 | $< 0.001^{**}$ |
| $\leq 5$ MW  | 0.000                                 | 0.000           | 0.932          | -                                      | -              | -          | 0.341                               | 0.293<br>0.398 | $< 0.001^{**}$ |
| Absent   | 2.078                                 | 1.559<br>2.772  | $< 0.001^{**}$ | -                                      | -              | -          | 0.467                               | 0.418<br>0.521 | $< 0.001^{**}$ |
| <b>Employment</b>                                    |                                       |                 |                |  |                |            |                                     |                |                |
| Formal   | -                                     | -               | -              | -                                      | -              | -          | 1.00*                               | -              | -              |
| Informal   | -                                     | -               | -              | -                                      | -              | -          | 1.641                               | 1.447<br>1.860 | $< 0.001^{**}$ |
| Unemployed and Refusal                               | -                                     | -               | -              | -                                      | -              | -          | 3.164                               | 2.866<br>3.493 | $< 0.001^{**}$ |
| <b>Built environment (average z-scores per zone)</b> |                                       |                 |                |  |                |            |                                     |                |                |
| Bike   | 0.402                                 | 0.148<br>1.096  | 0.075          | -                                      | -              | -          | -                                   | -              | -              |
| Building   | -                                     | -               | -              | 0.789                                  | 0.541<br>1.151 | 0.219      | -                                   | -              | -              |
| Bus  | -                                     | -               | -              | -                                      | -              | -          | -                                   | -              | -              |
| Car  | -                                     | -               | -              | -                                      | -              | -          | -                                   | -              | -              |
| Fence  | -                                     | -               | -              | -                                      | -              | -          | -                                   | -              | -              |
| Motorcycle Person                                    | -                                     | -               | -              | -                                      | -              | -          | 1.857                               | 1.789<br>1.931 | $< 0.001^{**}$ |
| Pole   | 0.333                                 | 0.153<br>0.724  | $< 0.001^{**}$ | -                                      | -              | -          | -                                   | -              | -              |
| Rider  | 8.361                                 | 2.274<br>30.743 | $< 0.001^{**}$ | -                                      | -              | -          | -                                   | -              | -              |
| Road   | -                                     | -               | -              | -                                      | -              | -          | 0.317                               | 0.221<br>0.456 | $< 0.001^{**}$ |
| Sidewalk   | -                                     | -               | -              | -                                      | -              | -          | 4.999                               | 3.986<br>6.269 | $< 0.001^{**}$ |
| Sky  | -                                     | -               | -              | -                                      | -              | -          | -                                   | -              | -              |
| Terrain  | -                                     | -               | -              | 0.476                                  | 0.288<br>0.784 | 0.003**    | -                                   | -              | -              |
| Train  | -                                     | -               | -              | -                                      | -              | -          | -                                   | -              | -              |
| `T. Light`   | -                                     | -               | -              | -                                      | -              | -          | -                                   | -              | -              |
| `T. Sign`  | -                                     | -               | -              | -                                      | -              | -          | -                                   | -              | -              |
| Truck  | -                                     | -               | -              | 0.591                                  | 0.336<br>1.037 | 0.06.      | -                                   | -              | -              |
| Vegetation   | -                                     | -               | -              | 0.641                                  | 0.445<br>0.923 | 0.016*     | 0.344                               | 0.268<br>0.443 | $< 0.001^{**}$ |
| Wall   | 2.216                                 | 1.259<br>3.899  | $< 0.001^{**}$ | -                                      | -              | -          | -                                   | -              | -              |
| Integration I200                                     | 1.00                                  | 0.998<br>1.010  | 0.274          | 3.732                                  | 0.858<br>16.22 | 0.078      | -                                   | -              | -              |
| Inters.Density                                       | -                                     | -               | -              | -                                      | -              | -          | -                                   | -              | -              |
| Res. Density   | -                                     | -               | -              | 0.890                                  | 0.747<br>1.059 | 0.189      | -                                   | -              | -              |
| Retail FAR   | -                                     | -               | -              | -                                      | -              | -          | -                                   | -              | -              |
| Land Use Mix   | -                                     | -               | -              | -                                      | -              | -          | 1.452                               | 1.252<br>1.685 | $< 0.001^{**}$ |
| <b>Model Fit</b>                                     | <b>AIC = 1783.5<br/>logLik=-877.7</b> |                 |                | <b>AIC = 3935.5<br/>logLik=-1958.7</b> |                |            | <b>AIC = 17114<br/>logLik=-8542</b> |                |                |

Note<sup>1</sup>: \*=1,00: reference category; CI 95%: 95% Confidence Interval. \*\* significative at the  $\alpha=0,05$  level.

Source: The authors. 2023.

The primary objective of the study was **to elucidate multi-scale walkability in medium-small Brazilian towns**. This was accomplished by focusing on the creation of neighborhood-level walkability models considering micro and macro scale BE feature that best predict walking (vs not). A comprehensive examination of urban form and pedestrian behavior formed the basis of these models. In parallel, deep learning and statistical methodologies were explored. These novel tools enhanced the process of data collection and model construction in the analysis of walkability. This aim involved an evaluation of the benefits derived from the integration of micro-scale factors into macro-scale analyses, with their interplay with walking behavior taken into account. Finally, the study sought to examine contextual urban form, with attention given to unique local spatial nuances.

The conducted multi-scale assessment of walkability, in both micro and macro scales, is fundamentally based on a comprehensive examination of urban form and pedestrian behavior. These models were provided unique contextual insights into walkability dynamic for each case study (

Table 19). A deep learning computer vision methodology was explored to supplement the traditional methods of data analysis. This technique, which represents the frontier of urban data science, served as powerful tools in the process of data collection and model construction for walkability analysis. This application resulted in more refined models, capable of capturing complex relationships and subtle patterns that might otherwise remain unnoticed in macro scale city wide analysis. This approach enables a more scalable and efficient assessment of micro-scale walkability features, with the ability to process large amounts of visual data quickly and accurately, this approach can dramatically improve the granularity and scale of walkability assessments in low- and middle-income countries' small and average towns.

For public urban planning municipal bodies, especially in low and middle-income countries, these methodologies open new possibilities. They enable the systematic monitoring of the built environment and facilitate the achievement of longitudinal targets for urban planning and design improvements. This is particularly significant for medium and smaller-sized towns, where resources for such extensive analyses might be limited. By harnessing the capabilities of deep learning and computer vision, these towns can effectively assess and monitor walkability, leading to more informed decision-making. This, in turn, has the potential to contribute to more walkable, sustainable, and equitable urban environments.

This study sought **to evaluate the benefits derived from the integration of micro-scale factors into macro-scale analyses**, with their interplay with walking behavior

considered. Observations from Rolândia revealed that the presence of Poles, that includes sign pole, traffic light poles for example, were related to walking with ORs of less than 1, implying that an increase in prevalence of these elements decreases the probability of a trip being undertaken by walking. Contextual factors, such as Rolândia being intersected by a highway, can be conjectured to be attributed to these findings. Where individual might not walk in these busy car-traffic areas and concentrate walking patterns intra-neighborhoods. “Urban highways” as a built environment feature have been considered significant barrier to walking for residents along with general car-oriented areas, with a high presence of heavy vehicles (Lucchesi et al., 2023). Similarly, the presence of walls emerged as a positive determinant for walking, with ORs exceeding two. This suggests a possibility of individuals choosing walking over other means of transport in neighborhoods dominated by residential typologies, often marked by walls and fences in the broader Brazilian context.

Furthermore, the presence of individuals riding motorcycles or bicycles was recognized as a positive determinant for walking in Rolândia, with high OR exceeding 8. Rolândia is among the Brazilian cities with the highest levels of bicycle commuting, where 19.6% of all trips are made by bicycle, which elevates the city of Rolândia to a European level of bicycle use as a mode of transport (Fernandes et al., 2021). Thus, it can be inferred that areas with a high prevalence of this active mode of transportation might also exhibit substantial pedestrian traffic. Importantly, none of the macro variables demonstrated significant importance in promoting walking within this sample for the city of Rolândia. This could imply a greater significance of micro-scale characteristics in determining walking behaviors in a small city, as has been explored by (de Castro & Kanashiro, 2021).

In the Cambé model, two micro-scale variables were identified as relevant for pedestrian travel. These variables exhibited ORs of less than 1, suggesting a decrease in the likelihood of opting to walk as these variables increase. The presence of Terrain, for example vacant lots, and vegetation was found not to stimulate active commuting by foot. The presence of terrain can be related to perceptions of unsafety and incivilities, therefore this result can be corroborated by literature (Y. Kang et al., 2023).

For both Cambé and Londrina, the result for the Vegetation category contradicts literature references from high income countries where vegetation is a widely recognized positive influence on walking (Dadvand et al., 2016; Long & Liu, 2017b; Lu, 2018b) despite possible discussions otherwise (Groff & McCord, 2012), and general assumptions regarding the Brazilian climate. Green streets are understood to facilitate walking behaviors by making

walking routes attractive and comfortable, reducing heat, noise, and air pollution (Tilt et al., 2007). Further, considering correlate studies that apply machine learning studies in high and upper middle-income countries to urban analysis and walking (LONG; LIU, 2017; LU, 2018a), this result is also adverse. However, evidence provided by previous studies regarding the association between urban greenness and physical activity in LMICs and Brazil specifically, indicates otherwise, that large presence of vegetation, are considered detrimental to the walkability perception and consequently to the walking frequency (Lucchesi et al., 2023).

Considering the current production of the theme, this study seems to be the first to quantitatively assess eye-level greenness through GSV for entire cities in small and medium cities from a LMIC. Therefore, this result can be related to social aspects not yet explored. Contextually, those living in areas with more vacant lots and more vegetation tend to walk less, implying potential housing options in peripheral areas and specifically for this case study, the agglomerated areas with the city of Londrina, abundant in vegetation and potential vacant plots, which might dissuade walking. It's crucial to emphasize that while proximity to nature is desirable, it demands certain infrastructural prerequisites. Large vegetative areas, while possibly aesthetically pleasing, necessitate infrastructure to prevent them from becoming unsafe. Furthermore, the three cities selected as case studies are planned cities (Rego & Meneguetti, 2006), featuring orthogonal layouts with deliberate proportions of streets, sidewalks, and green spaces. This inherent design aspect might significantly influence the findings.

Regarding the built environment results for the city of Londrina, this was the only city where a macro scale variable proved relevant. Mixed land use influenced the choice to walk with an OR of 1.45, suggesting that an increase in mixed land use heightens the likelihood of an individual choosing to walk. Notably, in areas with more car-focused paved roads, the chances of walking decrease, while the presence of more sidewalks increases the likelihood of walking, it can be considered that these results align with current evidence on sidewalk provision (Adams et al., 2022). In terms of urban design, this result is relevant as it reiterates the importance of sidewalks not only as transportation facilities but also support for stationary activities, serving as infrastructure, spaces of everyday life, and eventually as leisure destinations (Ehrenfeucht & Loukaitou-Sideris, 2010). Increasing sidewalk availability and quality should be a recurring concern of urban developments. It scientifically demonstrates itself as efficient use of resources that is cost-effective and effectively encourages walking (Gunn et al., 2014).

In Londrina, the odds of a trip being walked higher for an individual residing in a zone with higher values of average z-score for the person category, that is, having more people on the streets the outcome is more likely to occur. The presence of people and their movement can indicate the quality of the micro-scaled BE, where this presence indicates and interaction and available opportunities for such social phenomenon to occur. To analyze and create walkable communities, assessing pedestrian volumes frequently is necessary (R. Ewing & Clemente, 2013). However, it remains unclear whether machine learning techniques can reliably assess pedestrian volume (L. Chen et al., 2020). The evidence found here builds on the possibility of using the methodology applied here for pedestrian activity identification in the field of walkability and urban planning. This representation of people's presence might be even more strongly related to Vitality, an urban quality more connected to the presence (J. Jacobs, 1961). rather than to people's movement as Walkability. However, these conjectures should be analyzed in more focused studies.

Upon analyzing the results for the cities of Rolândia, Cambé, and Londrina, within the context of the "Aspects of Behavior-Environment" framework by Moudon and Lee (2003), relationships emerge between the different contexts and their respective relevant walkability elements. In Rolândia, the smallest city with a population of approximately 60,000, the walkability elements were primarily observed within the Spatiopsychosocial and Spatiobehavioral categories. The inclusion "poles," aligns with the Spatiopsychosocial category, under the facet of "Lighting." Additionally, traffic signaling finds alignment within the Spatiobehavioral domain, particularly resonating with "Street supports for walking." The variable "rider," denoting individuals on bikes or motorcycles, aligns with the Spatiobehavioral category due to "Vehicle-pedestrian interactions." Furthermore, the element "walls" might be classified under Spatiopsychosocial, especially when influencing factors such as "View/surveillance" or "Safety." The socio-psychological aspects, such as community ambiance and aesthetic values, appear to be influential in the smallest case study, perhaps due to the limited area and compactness of the city which allows for more exposure to social urban life.

Cambé, a moderately populated city with an estimated 100,000 residents, presented elements that resonated with the Spatiophysical and Spatiopsychosocial categories. Specifically, the "terrain" variable associates with the Spatiophysical domain, emphasizing "Walking path/sidewalks" attributes. Concurrently, the "vegetation" variable is pertinent to both the Spatiophysical category, in the context of green spaces, and the Spatiopsychosocial sphere,

in terms of its potential impact on aesthetics. The city's growth seems to reflect an increasing emphasis on both tangible environmental features and socio-psychological needs, suggesting a transitional urban development phase, possibly also related to the urban agglomeration process with Londrina.

Londrina, being the largest with nearly 600,000 inhabitants, revealed a complex micro scale walkability landscape, covering the Spatiobehavioral, Spatiopsychosocial, and Spatiophysical categories. The variable "person" hints at implications for the Spatiobehavioral category, whereas infrastructure elements like "road" and "sidewalk" converge with the Spatiophysical category. Simultaneously, the "vegetation" element echoes observations from Cambé, interacting with the Spatiophysical domain and the Spatiopsychosocial category. The city's dense population and vast urban expanse seem to result in a walking environment that acts upon individuals in a complex way through nuanced interactions between inhabitants and their environment.

In summary, the comparative insights elucidate that the walkability determinants of these cities, relative to their sizes, vary considerably. While smaller cities emphasize socio-psychological and behavioral aspects, larger urban environments reflect a more holistic approach that integrates various determinants.

#### **4.5 Evaluation of the design process and discussions of multiscale walkability**

The results from multilevel modeling in three distinct cities showcased diverse patterns. In Rolândia, the smallest city studied, it was found that factors such as the presence of walls and individuals on bikes or motorcycles positively influenced the decision to walk. However, the existence of urban signage like posts and traffic lights had the opposite effect, reducing pedestrian movement. In Cambé, a medium-sized city with around 100,000 residents, the presence of vacant lots and green spaces was associated with a reduction in walking. For the largest city in the study, an array of micro and macro variables impacted the likelihood of walking. Increased sidewalk availability and mixed land use positively correlated with walking. Still, a higher presence of roads and greenery was negatively associated with walking, implying that as these variables increase, the probability of walking decreases.

These findings highlight that city size can influence walking behaviors and that urban factors affecting these behaviors can vary significantly. Rolândia's results suggested that urban safety amenities, often associated with heavy traffic, discouraged walking, while neighborhood-specific factors were more influential. Cambé, being intensively intertwined with the city of

Londrina, provided less clear insights, suggesting that future studies should focus on similar agglomerated city contexts for further comprehension by the walkability field of study of this phenomenon. For Londrina, the findings suggested that land use mix as an essential larger scale variable, as well as basic infrastructure such as sidewalks and the social aspect of people's presence, become increasingly relevant as city size expands. This can also be further explored in future studies that can add upon this work to include larger cities from Brazil.

Considering these results, it becomes clear that the macro-scale is less influential in determining the decision to walk in small to medium-sized towns in Brazil when compared to the micro-scale. The Hierarchy of Walking Needs Hypothesis by Alfonzo (2005), outlines that the fundamental prerequisite for walking is feasibility. This is influenced by factors such as age, physical condition, and available time. Once feasibility is established, accessibility becomes a priority, followed by higher-level needs like safety, comfort, and overall walking experience quality. Research indicates that accessibility is largely driven by broad neighborhood-level factors, whereas the higher-order needs are closely tied to specific street-level details. For example, while accessibility examines factors like distance and walking infrastructure, the higher-order needs assess streetscape characteristics, sidewalk widths, and the inclusion of amenities like street trees or benches. Alfonzo's hypothesis emphasizes that individuals won't prioritize higher-level walking needs if their foundational needs aren't satisfied. This study shows that although large-scale factors are typically more influential in walking behavior for larger sized cities, in smaller one, the micro-scale elements also have a significant impact perhaps more substantial than that of accessibility.

In terms of walking data, initially a critical comparative assessment of average walking distances can be constructed. Upon comparing the average walking distances in each city, distinct variations emerge: Londrina has an average of 1.169 km, Rolândia registers a slightly lower 1.116 km, and Cambé stands out with a higher 1.463 km. These discrepancies not only highlight divergent pedestrian behaviors but also prompt a critical examination of the applicability of contemporary urban theories in these contexts, especially for smaller towns such as Rolândia. For instance, when juxtaposing these findings with the concept of the "15-minute city" (Moreno et al., 2021), an alternative planning approach to improve quality of life in cities with the 15-minute walkable neighborhood, it becomes evident that such modern trends may not always align with localized realities. Similarly, to the results obtained in this study for the smaller cities that contradict the assumptions based on hierarchies of needs, these contemporary urban planning theories might not be universally suitable. Consider the spatial

layout of Rolândia: traversing the city entirely from east to west covers approximately 8 km, and from north to south, it spans around 6 km. Yet, the average utilitarian walking distance in the city stands at just 1.116 km - roughly a mere 15-minute walk. This striking contrast accentuates the need for more tailored urban planning strategies that account for the unique characteristics and needs of individual cities, rather than relying solely on generalized contemporary concepts.

An argument can be made regarding the predominant planning characteristics across cities, irrespective of their size, by considering the prominent features present throughout the three cities. These findings are summarized in Table 20 provide insights on the general composition of streetscapes with indications on the percentage of each class, as identified through the semantic segmentation procedure across all case studies. The results show that the categories of Sky, Road, Vegetation, and Building are predominant in the streetscape. Notably, the Sky category emerges as the most significant class across all cities. This could indicate that cities, on the whole, possess low levels of enclosure—a feature that tends to foster feelings of safety for pedestrians (R. Ewing & Handy, 2009). The substantial presence of the sky could be attributed to widespread urban sprawl and a prevalence of horizontal buildings, expansive peripheries near open fields, and edges dotted with vacant lots. The Road category is also significant, reflecting a modern 'rodoviarist' culture that lacks adequate transport infrastructure to support modes of travel other than private cars (H. Badland et al., 2014). Vegetation is another dominant feature. As previously discussed, vegetation can potentially enhance walkability by offering supportive environments and reducing exposure to air pollution, heat, and noise (Dadvand et al., 2016; Lu, 2018b). However, based on the results from predictive models, it's evident that, at least for Cambé and Londrina, this characteristic has a complex relationship with walking behaviors, potentially deterring rather than promoting it.

Table 20 – Percentage classes from all obtained GSV images for all cities.

| Classes       | Londrina | Cambé | Rolândia |
|---------------|----------|-------|----------|
|               | %        | %     | %        |
| Sky           | 33.57    | 39.19 | 35.21    |
| Road          | 26.66    | 28.67 | 28.15    |
| Vegetation    | 14.48    | 10.46 | 14.24    |
| Building      | 5.94     | 4.92  | 4.55     |
| Terrain       | 7.10     | 7.86  | 7.17     |
| Sidewalk      | 5.01     | 4.26  | 4.57     |
| Fence         | 2.60     | 1.97  | 2.32     |
| Car           | 1.53     | 0.88  | 0.71     |
| Wall          | 2.23     | 0.19  | 0.20     |
| Pole          | 0.53     | 0.58  | 0.60     |
| Person        | 0.05     | 0.05  | 0.05     |
| Truck         | 0.13     | 0.18  | 0.09     |
| Bicycle       | 0.02     | 0.02  | 0.03     |
| Bus           | 0.03     | 0.01  | 0.01     |
| Rider         | 0.02     | 0.01  | 0.02     |
| Traffic Sign  | 0.03     | 0.03  | 0.03     |
| Train         | 0.04     | 0.03  | 0.02     |
| Motorcycle    | 0.02     | 0.01  | 0.01     |
| Traffic Light | 0.00     | 0.00  | 0.00     |

Source: The author, 2023

Overall, this study reveals that in a Brazilian context both broad and detailed environmental factors play a role in influencing walking behavior, however this synergistic relationship is more relevant as city size increases. This understanding is unprecedented, as until the time of this study, no comprehensive analyses encompassing both the micro and macro scales across entire cities have been conducted, to the best of the authors' knowledge.

Acknowledging walkability as a complex construct that invariably depends on social, cultural, and economic factors (Shashank & Schuurman, 2019), this initial finding signifies a broad field of study within the Brazilian context. Furthermore, it is essential to understand that these models do not necessarily reflect preferred areas for walking. Rather, they indicate where individuals in these contexts are walking and expose the circumstances, they encounter within the urban environments they traverse. In the obtained models, individuals with lower incomes and, for the case of Londrina, those unemployed or informally employed, were more likely to choose walking as their mode of transportation. In LMICs especially, such individuals are frequently exposed to urban environments with conditions that are less than ideal and therefore

might not be engaging in physical activity that is health-enhancing (Salvo et al., 2023), further emphasizing the need for social justice in urban planning.

Further, in face of the obtained results and the contextual urban form, spatial, social, and cultural nuance, the importance of considering spatial inequalities within Brazilian cities must be emphasized. To further incorporate justice in urban policies, minimize health disparities, and improve access to high-quality urban environments, in terms of walkability, the spatial inequities within cities must be addressed. These inequities can often stem from inconsistent urban planning policies that have led to unequal distribution of walkable spaces and public amenities. This directly impacts the health and quality of life of residents in less privileged areas. To address these disparities, an approach of distributive justice despite vital is not enough, going beyond fair allocation of walkability features, considering factors such as safety, aesthetics, access to amenities, and the presence of sidewalks and crosswalks. It's essential to consider the processes that lead to these unjust distributions and that participatory planning is a crucial element of urban planning justice. Moreover, existing policies and rights, often viewed as standard today, stem from historical events and political debates characterized by power imbalances among social groups. Understanding how society views the role and essence of walkable environments is paramount, as it determines the perception of what just urban planning policies targeted to active transportation is. The idea that justice involves more than just distribution requires placing distributive justice within a broader frame, encompassing participatory planning, democratic citizenship, urban rights, and spatial justice (Pereira et al., 2017; Pereira & Karner, 2021).

Accessing walkable environments is influenced by how individuals interact with their surroundings. Factors like gender, age, disabilities, and time constraints contribute to the varying accessibility experiences among people (Pereira et al., 2017). Considering gender disparities, the limited scholarly focus on women's experiences as pedestrians has led to an apparent research gap in understanding women's transportation patterns, as noted by Golan et al. (2019). Existing studies, largely devoted to analyzing these patterns within urban characteristics, typically highlight substantial inequities between men and women. These disparities encompass a range of aspects, including perceptions of urban landscape attributes, fear of crime, and the influence of land use patterns on their walking levels (Clifton et al., 2005). Women's fear of crime has also been emphasized as a significant factor shaping their mobility patterns (Koskela & Pain, 2000; R. H. Pain, 1997). This body of research, while illuminating,

is still expanding, with recent contributions in the Brazilian contexts (Leão et al., 2021; Leão & Kanashiro, 2021)

Beyond focusing solely on gender, it's essential to integrate an intersectional perspective that encompasses both gender and race. There are significant challenges in cultivating a culture that holistically considers the intersection of gender and race when planning and executing public mobility policies. Historically, there has been a notable lack of diversity in transport and urban planning sectors, particularly in upper decision-making roles (ITDP, 2021). Most Origin-Destination surveys, the main tools used to understand city movement patterns, do not include race data. This omission could lead to public policies that perpetuate existing inequalities. Emphasizing that gender and race need joint analysis is vital to truly grasp exclusionary experiences, particularly those faced by black women. Brazilian cities must emphasize increasing diversity in technical bodies for urban planning and data collection on travel patterns that responds to both gender and race, championing diversity in urban mobility roles across gender, race, class, age, and disabilities.

The issue of database composition in the context of urban planning and management, particularly regarding mobility plans, is worth discussing. In response to the pressing need for a more sustainable and environmentally balanced urban mobility system, the Brazilian Federal Government enacted Law 12.587 in 2012, establishing the National Urban Mobility Policy (PNMU). As an urban planning tool, its primary objective is to enhance the integration of various transportation modes and improve accessibility and mobility for both people and goods. This legislation, primarily aimed at fostering sustainable urban development, provides Brazilian municipalities with tools and guidance to enhance mobility conditions, with a clear emphasis on the pedestrian and active transportation as priority. Its guiding principles include universal accessibility, sustainable city development in both socio-economic and environmental dimensions, equal access to public transportation, and efficiency in urban transportation services. However, while these guidelines exist, there's a noticeable lack of transparency and uniformity at the federal level.

Work conducted from 2021-2023 sourcing data from the National Secretary of Mobility and Urban and Regional Development in September 2022<sup>4</sup> reveals that out of the 386 reported mobility plans to exist in Brazil, only 201 are publicly accessible for review. This indicates a

---

<sup>4</sup> Data reference: National Secretary of Mobility and Urban and Regional Development (September 2022). As cited in the technical report from the CNPQ Universal project titled "Active mobility, health and wellbeing: multiscalar assessment of walkability in BR cities". Process number: 422644/2021-8. Originating from the call: CNPq/MCTI/FNDCT No 18/2021 - Faixa A - Grupos Emergentes.

significant opacity and lack of consistent federal oversight in the country's urban mobility processes. Furthermore, while there are existing guidelines on the preparation of plans and reference materials available, there seems to be no systematic oversight or regulation regarding the alignment of these plans or even the methodology used in Origin-Destination surveys. Such inconsistencies and lack of clear federal oversight jeopardize the very essence of the PNMU, its objectives and pose negative implications for the monitoring of active behaviors. These disparities could affect the accuracy of data and, subsequently, the effectiveness of the mobility plans.

Further compounding these challenges is the lack of basic information, such as those related to land use. Quality data in this aspect is often unavailable within municipal administrations, let alone in a democratized and accessible format. Without tracking this type of information, it is nearly impossible to efficiently monitor and create zoning plans and that effectively address longitudinal changes. Basic data on the spatial distribution of environmental characteristics in cities and the behaviors of individuals are crucial for robust planning and management.

Open data can be seen as a solution for this issue, however stark data inequalities persist within open databases such as OpenStreetMap (OSM). This spatial bias is distributed across various Human Development Index groups, population sizes, and geographic regions (Herfort et al., 2021), including medium and small Brazilian cities. These disparities have concrete impacts, particularly in the realm of walkability analysis in LMICs, a key aspect of urban design and public health. In these settings, the availability of open data would be advantageous due to the limited technical support available within city governing bodies and agencies, and the chronic lack of funding for maintaining such data over extended periods. Thus, the persistent disparities in data availability and quality pose significant challenges to effective research and practice of urban planning and equitable development. Implementing such open data could likely result in more comprehensive comparisons of factors like entropy for multiple contexts.

In addition, it must be highlighted that relying solely on open data and satellite imagery as solutions for measurement of urban qualities must be regarded with caution, as in LMICs especially, spatial inequities play a major role in accessing amenities and built environment factors that are recognized as health enhancing and positive for urban life (Pereira et al., 2017). Existing structural inequities dictate the accessibility to walkability and liveability features, for example, therefore spatial inequalities must not be overlooked, at risk of an oversimplification of the complex social phenomenon that is urban living.

Using the Design Science Research framework, this section clarified two central methodological steps. Initially, there was an Evaluation phase during which the walkability models, termed here as artifacts, were critically assessed. This is followed by the Conclusion phase once the research results were determined to be successful. A comprehensive assessment showed that the artifacts, specifically the developed models, functioned effectively within their intended methodological framework. Furthermore, the thesis, which suggested the potential of a multiscale neighborhood-level walkability analysis for medium-small Brazilian towns using deep learning, was strongly supported.

When viewed in broader terms, the microscale was analyzed along with the macro-level, shedding light on the environments that individuals walk in medium and small Brazilian towns. Even though the core capabilities of the tool are recognized, the focus was primarily on the broader implications of its outcomes, especially concerning urban management and the conceptualization of walkability studies in Brazil. The findings highlighted the imperative to weave discussions on structural social disparities into the larger walkability narrative in Brazil, factoring in the distinct attributes of smaller towns and the wide-ranging elements affecting walkability. The groundwork established by this research aimed to catalyze more in-depth examinations and understandings of the complex interplay between urban planning, societal dynamics, and walkability.

The research presented in this thesis advances the field by automating a traditionally costly and largely non-existent surveying process of micro-scaled features of the BE in planning institutes and municipal administrations of smaller Brazilian cities. This automation can contribute to more efficient urban management, especially in relation to the micro-scale. However, it is essential to underscore that the understanding and monitoring of macro-scale characteristics are equally important within the context of urban planning and management.

Despite potential restrictions on the public usage of Google Street View (GSV), this approach's adaptability holds promise. With further enhancements and the development of a user-friendly interface, this approach could become an invaluable tool for city planners in LMICs. It would allow them to monitor city developments effortlessly and with a high degree of traceability, combining the power of advanced tools for the development of more evidence based rather than intuitive planning. This multi-scalar approach, considering both micro and macro aspects, along with spatial inequity considerations can ensure more holistic urban management and planning, bringing us one step closer to achieving equitable and walkable Brazilian cities.

## 5 CONCLUSIONS

The multilevel modeling results for the three cities revealed distinct trends. In the smallest city, Rolândia, the presence of walls and individuals riding bicycles or motorcycles increased the odds of a journey being undertaken by walking. Conversely, the presence of signaling elements such as posts and traffic lights decreased the likelihood of a journey being taken on foot. In the city of Cambé, which has approximately 100,000 inhabitants and represents the medium-sized city in this study, it was found that the presence of vacant lots and vegetation reduced the odds of a journey being taken on foot. Finally, for the largest city in the study, a mix of macro and micro variables proved relevant for predicting the choice to walk. The presence of sidewalks and mixed land use had the ability to increase the odds of a journey being walked, indicated by an OR greater than 1. This suggests that the greater the existence of these elements, the higher the likelihood of a journey being walked. However, the presence of roads and vegetation was negatively correlated with the odds of walking. That is, the higher the levels of these variables, the lower the chances of a journey being walked.

These results underscore that the urban landscape of a smaller city possesses unique characteristics distinct from those of larger cities. For example, in the smallest city analyzed in this study, urban security amenities - often associated with high-traffic roads - are less walked by the local population. Small Brazilian cities potentially need the development of specific analyses and public policies tailored to these contexts. Different mobility patterns are observable in these smaller cities, which, when compared to traditional concepts like the '15-minute city,' reveal that conventional ideas of walkability may not apply due to the city's limited size. Therefore, there's a need for a more nuanced understanding of walkability in the context of smaller urban spaces. In contrast, elements that may be more related to intra-neighborhood areas proved more relevant. Cambé presents a very specific context, being intensively agglomerated with the city of Londrina. Hence, its results are slightly less interpretable and compromised by this fact. Future studies should aim to investigate urban agglomerations as an object of study. The city of Londrina suggested that as the city size increases, macro-scale variables might become more relevant. Moreover, variables related to basic infrastructures such as the presence of sidewalks, and the social element of the presence of people, become relevant as the city size increases.

Considering the thesis that *An efficient multiscale neighborhood-level walkability analysis for medium-small Brazilian Towns can be constructed through the use of deep learning*

*techniques.*, the contributions this research will provide to the walkability field of study include (a) first and foremost a general better understanding of the walkability phenomenon in the Brazilian town context, (b) the proposition of both a model for quantifying walkability and a novel method for gathering and analyzing walkability data and (c) the delivery of politically relevant evidence that has potential to guide urban planning strategies towards positively influencing active travel outcomes.

Technically robust and data-driven studies are necessary on the relationship between the built environment and active transportation, considering human behavior as a complex phenomenon that can't be precisely explained by simple models of exposures and individual characteristics (Dean et al., 2020b). In this proposal, we explicitly address this gap.

The literature emphasizes the need for policy-relevant interdisciplinary research that may lead to more contextually desirable outcomes (Sallis et al., 2016). This proposal goes towards this recommendation, providing a scholarly foundation for the development of walkable urban environments with an emphasis on local evidence. By using geoprocessing and cutting-edge statistical/computational tools, the proposed research has prospects of aggregating value in the development of evidence-centered, contextually tailored urban planning in the prospect of creating more sustainable cities: connected, denser, destination-full, and walkable.

Further, urban research usually focuses on explaining urban phenomena and exerts lesser efforts in developing normative theories of its possible advancements. In the context of planning practice, the characterization of the spatial form should go beyond just morphological descriptions to constitute an essential basis for planning (Talen, 2002). Considering this gap, the proposal presented here will also fill an essential gap of bringing empirical evidence to light that might contribute to such a future perspective. Despite such possible advancements, many clear limitations remain and are brought forward:

The process of aggregating walking data has the potential to homogenize results. This use of areal aggregations is often driven by practical considerations and the need to safeguard the confidentiality of individual population members (Martin, 1998). A notable challenge with most zonal geographies is the undefined relationship with underlying population characteristics, giving rise to the Modifiable Areal Unit Problem (MAUP) (Openshaw, 1984). In addition to the aggregation of spatial data, the categorization of individuals often occurs without differentiating pedestrians based on their ability (Lo 2009, 148) or race, presenting an oversimplification of pedestrian diversity and, consequently, walkability. The homogenous concept of the "citizen" (Kwon & Akar, 2022), particularly in a diverse country like Brazil, is

detrimental to mobility-focused public policies and walkability research. This limitation is noteworthy as the data employed in this study is secondary, derived from traditional origin-destination surveys. Such a construction approach can impose certain restrictions on the comprehensiveness and nuance of the data.

In this study, the potential effects of a broad spectrum of explanatory variables on the odds of walking were examined across multiple case studies, including sociodemographic characteristics. However, this kind of data is not available for all case studies or for most Brazilian cities. The omission of elements characterizing the responding population presents a substantial limitation, necessitating careful interpretation of results. Differentiating between utilitarian and leisure walking may reveal more precise relationships with the micro-scale. Yet, due to the current sample size in this study, no definitive conclusions could be drawn, likely due to under sampling. Future multiscale walkability research can delve deeper into these distinctions and relationships.

Prospective research could seek to augment open data from sources such as Open Street Map, thereby more accurately representing low- and middle-income countries and addressing significant existing data inequalities that vary considerably across countries (Herfort et al., 2022). Moreover, the prospect of maintaining longitudinal studies in walkability research not only lies in the use of open data but also in the efficient automation of calculations. Future research can rely on these approaches to advance the field within the Brazilian context.

## REFERENCES

- Adams, M. A., Phillips, C. B., Patel, A., & Middel, A. (2022). Training Computers to See the Built Environment Related to Physical Activity: Detection of Microscale Walkability Features Using Computer Vision. *International Journal of Environmental Research and Public Health*, *19*(8). <https://doi.org/10.3390/ijerph19084548>
- Adlakha, D., & Parra, D. C. (2020). Mind the gap : Gender differences in walkability , transportation and physical activity in urban India. *Journal of Transport & Health*, *18*, 1–10. <https://doi.org/10.1016/j.jth.2020.100875>
- Aghaabbasi, M., Moeinaddini, M., Shah, M. Z., Shekari, Z. A., & Kermani, M. A. (2017). Evaluating the capability of walkability audit tools. *Sustainable Cities and Society*, *37*, 475–484. <https://doi.org/10.1016/j.scs.2017.12.001>
- Aken, J. E. van. (2004). Management Research Based on the Paradigm of the Design Sciences: The Quest for Field-Tested and Grounded Technological Rules. *Journal of Management Studies*, *41*(2), 219–246. <https://doi.org/10.1111/j.1467-6486.2004.00430.x>
- Alexander, C., Ishikawa, S., Silverstain, M., Jacobson, M., Fiksdahl-King, I., & Angel, S. (1977). *A pattern language : towns, buildings, construction*. Oxford University Press, USA.
- Alfonzo, M. A. (2005). To walk or not to walk? The hierarchy of walking needs. *Environment and Behavior*, *37*(6), 808–836. <https://doi.org/10.1177/0013916504274016>
- Appleyard, D., Gerson, M. S., & Lintell, M. (1981). *Livable Streets*. University of California Press.
- Arellana, J., Saltarín, M., Larrañaga, A. M., Alvarez, V., Arellana, J., Saltarín, M., Larrañaga, A. M., Alvarez, V., Arellana, J., Saltarín, M., & Larrañaga, A. M. (2019). Urban walkability considering pedestrians ’ perceptions of the built environment : a 10-year review and a case study in a medium-sized city in Latin America built environment : a 10-year review and a case study in a. *Transport Reviews*, *0*(0), 1–21. <https://doi.org/10.1080/01441647.2019.1703842>
- Associação Nacional de Transportes Públicos - Antp. (2018). Sistema de Informações da Mobilidade Urbana: Relatório geral 2016. *ANTP*, *1*(1), 1–110. <http://files.antp.org.br/simob/simob-2016-v6.pdf>
- Augusto, P., Rodrigues, F., Melo, M. P. De, Assis, M. R. De, & Palma, A. (2017). Condições socioeconômicas e prática de atividades físicas em adultos e idosos : uma revisão sistemática. *Brazilian Journal of Physical Activity and Health*, *22*(3), 217–232. <https://doi.org/10.12820/rbafs.v.22n3p217-232>
- Austin, P. C., Goel, V., & Van Walraven, C. (2001). An introduction to multilevel regression models. *Canadian Journal of Public Health*, *92*(2), 150–154. <https://doi.org/10.1007/bf03404950>
- Badland, H. M., Duncan, M. J., Oliver, M., Duncan, J. S., & Mavoa, S. (2010). Examining commute routes: Applications of GIS and GPS technology. *Environmental Health and Preventive Medicine*, *15*(5), 327–330. <https://doi.org/10.1007/s12199-010-0138-1>
- Badland, H. M., Opit, S., Witten, K., Kearns, R. A., & Mavoa, S. (2010). Can Virtual Streetscape Audits Reliably Replace Physical Streetscape Audits ? *Journal of Urban Health*, *87*(6), 1007–1016. <https://doi.org/10.1007/s11524-010-9505-x>
- Badland, H., Mavoa, S., Villanueva, K., Roberts, R., Davern, M., & Giles-corti, B. (2014). The development of policy-relevant transport indicators to monitor health behaviours and outcomes. *Journal of Transport & Health*, 1–8. <https://doi.org/10.1016/j.jth.2014.07.005>

- Barton, J., Hine, R., & Pretty, J. (2009). The health benefits of walking in greenspaces of high natural and heritage value. *Journal of Integrative Environmental Sciences*, 6(4), 261–278. <https://doi.org/10.1080/19438150903378425>
- Bauman, A. E., Reis, R. S., Sallis, J. F., Wells, J. C., Loos, R. J. F., & Martin, B. W. (2012). Correlates of physical activity: Why are some people physically active and others not? *The Lancet*, 380(9838), 258–271. [https://doi.org/10.1016/S0140-6736\(12\)60735-1](https://doi.org/10.1016/S0140-6736(12)60735-1)
- Ben-joseph, E., Lee, J. S., Seoul, M., Cromley, E. K., Laden, F., & Troped, P. J. (2015). Virtual and Actual: Relative Accuracy of On-Site and Web-based Instruments in Auditing the Environment for Physical Activity. *Health Place*, 19, 138–150. <https://doi.org/10.1016/j.healthplace.2012.11.001>.Virtual
- Berrigan, D., Pickle, L. W., & Dill, J. (2010). Associations between street connectivity and active transportation. *International Journal of Health Geographics*, 9(1), 20. <https://doi.org/10.1186/1476-072X-9-20>
- Biljecki, F., & Ito, K. (2021). Street view imagery in urban analytics and GIS: A review. *LANDSCAPE AND URBAN PLANNING*, 215. <https://doi.org/10.1016/j.lurbplan.2021.104217>
- Blecic, I., Congiu, T., Fancello, G., & Trunfio, G. A. (2020). Planning and Design Support Tools for Walkability: A Guide for Urban Analysts. *SUSTAINABILITY*, 12(11). <https://doi.org/10.3390/su12114405>
- Boarnet, M. G., Day, K., Alfonzo, M., Forsyth, A., & Oakes, M. (2006). The Irvine-Minnesota inventory to measure built environments: Reliability tests. *American Journal of Preventive Medicine*, 30(2), 153-159.e43. <https://doi.org/10.1016/j.amepre.2005.09.018>
- Booth, K. M., Pinkston, M. M., & Poston, W. S. C. (2005). Obesity and the Built Environment. *Journal of the AMERICAN DIETETIC ASSOCIATION*, May(Supplement), s110–s117. <https://doi.org/10.1016/j.jada.2005.02.045>
- Boutayeb, A., & Boutayeb, S. (2005). The burden of non communicable diseases in developing countries. *International Journal for Equity in Health*, 4(2), 1–8. <https://doi.org/10.1186/1475-9276-4-2>
- Brownson, R. C., Hoehner, C. M., Brennan, L. K., Cook, R. A., Elliott, M. B., & McMullen, K. M. (2004). Reliability of 2 instruments for auditing the environment for physical activity. *Journal of Physical Activity and Health*, 1(3), 191–208. <https://doi.org/10.1123/jpah.1.3.191>
- Brownson, R. C., Hoehner, C. M., Day, K., Forsyth, A., & Sallis, J. F. (2009). Measuring the Built Environment for Physical Activity: State of the Science. *American Journal of Preventive Medicine*, 36(4 SUPPL.), S99-S123.e12. <https://doi.org/10.1016/j.amepre.2009.01.005>
- Buck, C., Pohlbeln, H., Huybrechts, I., Bourdeaudhuij, I. De, Pitsiladis, Y., Reisch, L., & Pigeot, I. (2011). Development and application of a moveability index to quantify possibilities for physical activity in the built environment of children. *Health & Place*, 17(6), 1191–1201. <https://doi.org/10.1016/j.healthplace.2011.08.011>
- Cain, K. L., Geremia, C. M., Conway, T. L., Frank, L. D., Chapman, J. E., Fox, E. H., Timperio, A., Veitch, J., Van Dyck, D., Verhoeven, H., Reis, R., Augusto, A., Cerin, E., Mellecker, R. R., Queralt, A., Molina-García, J., & Sallis, J. F. (2018). Development and reliability of a streetscape observation instrument for international use: MAPS-global. *International Journal of Behavioral Nutrition and Physical Activity*, 15(1). <https://doi.org/10.1186/s12966-018-0650-z>
- Cain, K. L., Millstein, R. A., Sallis, J. F., Conway, T. L., Gavand, K. A., Frank, L. D., Saelens, B. E., Geremia, C. M., Chapman, J., Marc, A., Glanz, K., & King, A. C. (2014). Contribution of Streetscape Audits to Explanation of Physical Activity in Four Age

- Groups Based on the Microscale Audit of Pedestrian Streetscapes (MAPS). *SOCIAL SCIENCE & MEDICINE*, 116, 82–92.  
<https://doi.org/10.1016/j.socscimed.2014.06.042>. Contribution
- Campoli, J. (2012). *Made for Walking: Density and Neighborhood Form* (2<sup>o</sup> ed). Lincoln Institute of Land Policy.
- Carmona, M., Tiesdell, S., Oc, T., & Heath, T. (2010). *Public Urban Spaces: The Dimensions of Urban Design* (First Edit). Architectural Press.
- Cauwenberg, J. Van, Bourdeaudhuij, I. De, Meester, F. De, Dyck, D. Van, Salmon, J., Clarys, P., & Deforche, B. (2011). Relationship between the physical environment and physical activity in older adults : A systematic review. *Health & Place*, 17(2), 458–469.  
<https://doi.org/10.1016/j.healthplace.2010.11.010>
- Cerin, E., Leslie, E., Toit, L. d., Owen, N., & Frank, L. D. (2007). Destinations that matter: Associations with walking for transport. *Health and Place*, 13(3), 713–724.  
<https://doi.org/10.1016/j.healthplace.2006.11.002>
- Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2(3), 199–219.  
[https://doi.org/10.1016/S1361-9209\(97\)00009-6](https://doi.org/10.1016/S1361-9209(97)00009-6)
- Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 834–848. <https://doi.org/10.1109/TPAMI.2017.2699184>
- Chen, L., Lu, Y., Sheng, Q., Ye, Y., Wang, R., & Liu, Y. (2020). Estimating pedestrian volume using Street View images: A large-scale validation test. *Computers, Environment and Urban Systems*, 81(101481), 1–11.  
<https://doi.org/10.1016/j.compenvurbsys.2020.101481>
- Christian, H. E., Bull, F. C., Middleton, N. J., Knuiaman, M. W., Divitini, M. L., Hooper, P., Amarasinghe, A., & Giles-Corti, B. (2011). How important is the land use mix measure in understanding walking behaviour? Results from the RESIDE study. *International Journal of Behavioral Nutrition and Physical Activity*, 8(55), 1–12.  
<https://doi.org/10.1186/1479-5868-8-55>
- Cinnamon, J., & Jahiu, L. (2021). Panoramic street-level imagery in data-driven urban research: A comprehensive global review of applications, techniques, and practical considerations. Em *ISPRS International Journal of Geo-Information* (Vol. 10, Número 7). MDPI AG. <https://doi.org/10.3390/ijgi10070471>
- Clarke, P., Ailshire, J., Melendez, R., Bader, M., & Morenoff, J. (2010). Using Google Earth to conduct a neighborhood audit : Reliability of a virtual audit instrument. *Health & Place*, 16(6), 1224–1229. <https://doi.org/10.1016/j.healthplace.2010.08.007>
- Clifton, K. J., Livi, A. D., & Park, C. (2005). Gender Differences in Walking Behavior, Attitudes About Walking, and Perceptions of the Environment in Three Maryland Communities. *Research on Women's Issues in Transportation*, Vol. 2. p. 79-88, 79–88.
- Clifton, K. J., Livi Smith, A. D., & Rodriguez, D. (2007). The development and testing of an audit for the pedestrian environment. *Landscape and Urban Planning*, 80(1–2), 95–110.  
<https://doi.org/10.1016/j.landurbplan.2006.06.008>
- Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., & Schiele, B. (2016). The Cityscapes Dataset for Semantic Urban Scene Understanding. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3213–3223.
- Cullen, C. (1961). *The Concise Townscape* (Paperback). The Architectural Press is an imprint of Elsevier.

- Dadvand, P., Bartoll, X., Basagana, X., Dalmau-Bueno, A., Martinez, D., Ambros, A., Cirach, M., Triguero-Mas, M., Gascon, M., Borrell, C., & Nieuwenhuijsen, M. J. (2016). Green spaces and General Health: Roles of mental health status, social support, and physical activity. *ENVIRONMENT INTERNATIONAL*, *91*, 161–167.  
<https://doi.org/10.1016/j.envint.2016.02.029>
- de Castro, C. C., & Kanashiro, M. (2021). Walkability in small towns: Collating objective variables and in situ observations. *Arquiteturarevista*, *17*(2), 277–295.  
<https://doi.org/10.4013/arq.2021.172.06>
- Dean, J., Biglieri, S., Drescher, M., Garnett, A., Glover, T., & Casello, J. (2020a). Health and Place Thinking relationally about built environments and walkability : A study of adult walking behavior in Waterloo , Ontario. *Health and Place*, *64*(April), 102352.  
<https://doi.org/10.1016/j.healthplace.2020.102352>
- Dean, J., Biglieri, S., Drescher, M., Garnett, A., Glover, T., & Casello, J. (2020b). Thinking relationally about built environments and walkability : A study of adult walking behavior in Waterloo , Ontario. *Health and Place*, *64*(June 2019), 102352.  
<https://doi.org/10.1016/j.healthplace.2020.102352>
- Delclòs-Alió, X., Rodríguez, D. A., Medina, C., Miranda, J. J., Avila-Palencia, I., Targa, F., Moran, M. R., Sarmiento, O. L., & Quistberg, D. A. (2022). Walking for transportation in large Latin American cities: walking-only trips and total walking events and their sociodemographic correlates. *Transport Reviews*, *42*(3), 296–317.  
<https://doi.org/10.1080/01441647.2021.1966552>
- Devlin, J., Chang, M.-W., Lee, K., Google, K. T., & Language, A. I. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Association for Computational Linguistics (Org.), *Proceedings of NAACL-HLT* (p. 4171–4186). <https://github.com/tensorflow/tensor2tensor>
- Doersch, C., Singh, S., Gupta, A., Sivic, J., Efros, A., Doersch, C., Singh, S., Gupta, A., Sivic, J., Efros, A., Makes, W., & Look, P. (2012). What Makes Paris Look like Paris ? *ACM Transactions on Graphics (SIGGRAPH 2012)*, *31*(4).  
<https://doi.org/10.1145/2185520.2185597>
- Doiron, D., Setton, E. M., Brook, J. R., Kestens, Y., McCormack, G. R., Winters, M., Shooshitari, M., Azami, S., & Fuller, D. (2022). Predicting walking-to-work using street-level imagery and deep learning in seven Canadian cities. *Scientific Reports*, *12*(1).  
<https://doi.org/10.1038/s41598-022-22630-1>
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. <http://arxiv.org/abs/2010.11929>
- Dyck, V., Cerin, E., Bourdeaudhuij, I. De, Salvo, D., Christiansen, L. B., Macfarlane, D., Owen, N., Mitas, J., Troelsen, J., Aguinaga-ontoso, I., Davey, R., Reis, R., Sarmiento, O. L., Scho, G., Conway, T. L., & Sallis, J. F. (2015). Moderating effects of age , gender and education on the associations of perceived neighborhood environment attributes with accelerometer-based physical activity : The IPEN adult study. *Health & Place*, *36*, 65–73. <https://doi.org/10.1016/j.healthplace.2015.09.007>
- Ehrenfeucht, R., & Loukaitou-Sideris, A. (2010). Planning urban sidewalks: Infrastructure, daily life and destinations. *Journal of Urban Design*, *15*(4), 459–471.  
<https://doi.org/10.1080/13574809.2010.502333>
- Ellis, G., Hunter, R., Tully, M. A., Donnelly, M., Kelleher, L., & Kee, F. (2015). Connectivity and physical activity : using footpath networks to measure the walkability of built environments. *Environment and Planning B: Planning and Design*, *42*(1), 1–22.  
<https://doi.org/10.1068/b140039p>

- Elshahat, S., Rorke, M. O., & Adlakha, D. (2020). Built environment correlates of physical activity in low- and middle-income countries : A systematic review. *PLOS ONE*, 1–19. <https://doi.org/10.1371/journal.pone.0230454>
- Emery, J., Crump, C., & Bors, P. (2003). Reliability and Validity of Two Instruments Suitability of Sidewalks and Roads. *American Journal of Health Promotion*, 18(1), 38–47.
- Ewing, R., & Cervero, R. (2010). Travel and the built environment. *Journal of the American Planning Association*, 76(3), 265–294. <https://doi.org/10.1080/01944361003766766>
- Ewing, R., & Clemente, O. (2013). *Measuring Urban Design: Metrics for Livable Places* (A. C. N. and R. Ewing, Org.; 1st ed). Island Press.
- Ewing, R., Clemente, O., Handy, S., Winston, E., & Brownson, R. (2006). Measurement Instrument for Urban Design Quantities Related to Walkability. Em *Robert Wood Johnson Foundation, Active Living Research Program (ALR)*.
- Ewing, R., & Handy, S. (2009). Measuring the Unmeasurable : Urban Design Qualities Related to Walkability. *Journal of Urban Design*, 14(1), 65–84. <https://doi.org/10.1080/13574800802451155>
- Ferenhof, H. A., & Fernandes, R. F. (2016). Desmistificando a revisão de literatura como base para redação científica: método SFF. *Revista ACB*, 21(3), 550–563. [https://www.researchgate.net/profile/Helio\\_Ferenhof/publication/311972168\\_Desmistificando\\_a\\_Revisao\\_de\\_Literatura\\_como\\_base\\_para\\_Redacao\\_Cientifica\\_Metodo\\_SSF/links/5866537908aebf17d39ab49d.pdf](https://www.researchgate.net/profile/Helio_Ferenhof/publication/311972168_Desmistificando_a_Revisao_de_Literatura_como_base_para_Redacao_Cientifica_Metodo_SSF/links/5866537908aebf17d39ab49d.pdf)
- Fernandes, D., Urbano, M. R., & Kanashiro, M. (2021). Routing for safer rides: A space syntax approach to predict bicycle collisions in a Brazilian city. Em *Urbe* (Vol. 13). Pontificia Universidade Catolica do Parana. <https://doi.org/10.1590/2175-3369.013.e20200106>
- Frank, L. (2000). Land Use and Transportation Interaction: Implications on Public Health and Quality of Life. *Journal of Planning Education and Research*, 20(1), 6–22. <https://doi.org/10.1177/073945600128992564>
- Frank, L. D., Andresen, M. A., & Schmid, T. L. (2004). Obesity relationships with community design, physical activity, and time spent in cars. *American Journal of Preventive Medicine*, 27(2), 87–96. <https://doi.org/10.1016/j.amepre.2004.04.011>
- Frank, L. D., Sallis, J. F., Saelens, B. E., Leary, L., Cain, K., Conway, T. L., & Hess, P. M. (2009). The development of a walkability index: application to the Neighborhood Quality of Life Study. *British journal of sports medicine*, 44(13), 924–933. <https://doi.org/10.1136/bjism.2009.058701>
- Frumkin, H., Frank, L., Jackson, and R., & Washington. (2004). Urban Sprawl and Public Health: Designing, Planning, and Building for Healthy Communities. Em *Environmental health perspectives* (Vol. 1, Número 1). Island Press.
- Fry, D., Mooney, S. J., & Lovasi, G. S. (2020). *Assessing Google Street View Image Availability in Latin American Cities*. 552–560.
- Gebel, K., Bauman, A. E., Sugiyama, T., & Owen, N. (2011). Mismatch between perceived and objectively assessed neighborhood walkability attributes: Prospective relationships with walking and weight gain. *Health and Place*, 17(2), 519–524. <https://doi.org/10.1016/j.healthplace.2010.12.008>
- Gebel, K., Bauman, A., & Owen, N. (2009). Correlates of Non-Concordance between Perceived and Objective Measures of Walkability. *Annals of Behavioral Medicine*, 37(2), 228–238. <https://doi.org/10.1007/s12160-009-9098-3>
- Gehl, J. (1986). “Soft edges” in residential streets. *Scandinavian Housing and Planning Research*, 3(2), 89–102. <https://doi.org/10.1080/02815738608730092>

- Gehl, J. (1987). *Life between buildings: using public space*. (2011 revis, Número 1). Van Nostrand Reinhold Company.
- Gehl, J. (2013). *Cidade para pessoas* (2<sup>a</sup>). Editora Perspectiva.
- Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models* (1<sup>a</sup> ed). Cambridge University Press.
- Gibbs, R. (2012). *Principles of Urban Retail Planning & Development* (1st edtio). Jon Wiley and Sons.
- Gilderbloom, J. I., Riggs, W. W., & Meares, W. L. (2015). Does walkability matter? An examination of walkability's impact on housing values, foreclosures and crime. *Cities*, 42(PA), 13–24. <https://doi.org/10.1016/j.cities.2014.08.001>
- Giles-corti, B., Moudon, A. V., Lowe, M., Cerin, E., Boeing, G., Frumkin, H., Salvo, D., & Foster, S. (2022). What next? Expanding our view of city planning and global health, and implementing and monitoring evidence- informed policy. *The Lancet Global Health*. [https://doi.org/10.1016/S2214-109X\(22\)00066-3](https://doi.org/10.1016/S2214-109X(22)00066-3)
- Glazier, R., Weyman, J., & Creatore, M. (2013). Development and Validation of an Urban Walkability Index for Toronto, Canada. *Torontohealthprofiles.Ca*, 1–21.
- Golan, Y., Henderson, J., Wilkinson, N. L., Weverka, A., Journal, S., Use, L., & Wilkinson, N. L. (2019). Gendered walkability : Building a daytime walkability index for women. *The Journal of Transport and Land Use*, 12(1), 501–526.
- Grasser, G. (2014). Walkability and Public Health: Development of GIS-based indicators of walkability for surveillance and planning purposes in the city of Graz. *Dissertation submitted for the Academic Degree of DOCTOR OF MEDICAL SCIENCE at the Medical University of Graz Institute for Social Medicine and Epidemiology*, 234.
- Grasser, G., Titze, S., & Strongegger, W. J. (2016). Are residents of high-walkable areas satisfied with their neighbourhood? *Journal of Public Health*, 24(6), 469–476. <https://doi.org/10.1007/s10389-016-0744-5>
- Groff, E., & McCord, E. S. (2012). The Role of Neighborhood Parks as Crime Generators. *Security Journal*, 25(1), 1–24.
- Gunn, L. D., Lee, Y., Geelhoed, E., Shiell, A., & Giles-Corti, B. (2014). The cost-effectiveness of installing sidewalks to increase levels of transport-walking and health. *PREVENTIVE MEDICINE*, 67, 322–329. <https://doi.org/10.1016/j.ypmed.2014.07.041>
- Gustafsson, J. (2017). Single case studies vs. multiple case studies: A comparative study. *Academy of Business, Engineering and Science Halmstad University, Sweden*, 1–15.
- Hajna, S., Dasgupta, K., Joseph, L., & Ross, N. A. (2014). A call for caution and transparency in the calculation of land use mix: Measurement bias in the estimation of associations between land use mix and physical activity. *HEALTH & PLACE*, 29, 79–83. <https://doi.org/10.1016/j.healthplace.2014.06.002>
- Hanibuchi, T. (2019). Virtual audits of streetscapes by crowdworkers. *Health & Place*, 59(August), 102203. <https://doi.org/10.1016/j.healthplace.2019.102203>
- Harvey, C., & Aultman-Hall, L. (2016). Measuring Urban Streetscapes for Livability: A Review of Approaches. *Professional Geographer*, 68(1), 149–158. <https://doi.org/10.1080/00330124.2015.1065546>
- Harvey, C. W. (2014). *Measuring Streetscape Design for Livability Using Spatial Data and Methods*.
- Helbich, M., Yao, Y., Liu, Y., Zhang, J., Liu, P., & Wang, R. (2019). Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in Beijing, China. *Environment International*, 126(January), 107–117. <https://doi.org/10.1016/j.envint.2019.02.013>
- Herfort, B., Lautenbach, S., Anderson, J., Zipf, A., Albuquerque, P. De, Herfort, B., Lautenbach, S., & Zipf, A. (2022). *Investigating the digital divide in OpenStreetMap* :

*spatio-temporal analysis of inequalities in global urban building completeness  
Investigating the digital divide in OpenStreetMap : spatio-temporal analysis of  
inequalities in global.*

- Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., & Zipf, A. (2021). The evolution of humanitarian mapping within the OpenStreetMap community. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-82404-z>
- Hilland, T. A., Bourke, M., Wiesner, G., Garcia, E., Parker, A. G., Pascoe, M., & Craike, M. (2020). Correlates of walking among disadvantaged groups : A systematic review. *Health and Place*, 63(October 2019), 102337. <https://doi.org/10.1016/j.healthplace.2020.102337>
- Hillier, B., & Hanson, J. (1984). The social logic of space. Em *Barlett School of Architecture and Planning. University College London*. (1° ed). Cambridge university press. <https://doi.org/http://dx.doi.org/10.1017/CBO9780511597237>
- Hillier, B., Penn, A., Hanson, J., Grajewski, T., & Xu, J. (1993). Natural movement: or, configuration and attraction in urban pedestrian movement. *Environment and Planning B: Planning and Design*, 20(1), 29–66. <https://doi.org/10.1068/b200029>
- IBGE. (2010). População residente, total, urbana total e urbana na sede municipal - Paraná. Em *Censo 2010*.
- IBGE. (2018). *Panorâma da cidade de Rolândia*. <https://cidades.ibge.gov.br/brasil/pr/rolandia/panorama>
- IPARDES. (2018). *CADERNO ESTATÍSTICO MUNICÍPIO- MUNICÍPIO DE ROLÂNDIA* (Vol. 1, Número 1, p. 1–43).
- Jacobs, A. B. (1993). Great Streets. *ACCESS Magazine*.
- Jacobs, J. (1961). *The Death and Life of Great American Cities* (1ª Edição). VINTAGE BOOKS- Random House.
- Kang, B., Moudon, A. V., Hurvitz, P. M., & Saelens, B. E. (2017). Differences in behavior, time, location, and built environment between objectively measured utilitarian and recreational walking. *Transportation Research Part D: Transport and Environment*, 57, 185–194. <https://doi.org/10.1016/j.trd.2017.09.026>
- Kang, Y., Kim, J., Park, J., & Lee, J. (2023). Assessment of Perceived and Physical Walkability Using Street View Images and Deep Learning Technology. *ISPRS International Journal of Geo-Information*, 12(5). <https://doi.org/10.3390/ijgi12050186>
- Kim, S., Park, S., & Lee, J. S. (2014). Meso- or micro-scale? Environmental factors influencing pedestrian satisfaction. *Transportation Research Part D: Transport and Environment*, 30, 10–20. <https://doi.org/10.1016/j.trd.2014.05.005>
- Knell, G., Durand, C. P., Shuval, K., Kohl, H. W., Salvo, D., Olyuomi, A., & Pettee, K. (2019). If You Build It, Will They Come? A Quasi-experiment of Sidewalk Improvements and Physical Activity. *Translational Journal of the American College of Sports Medicine (TJACSM)*, 3(9), 66–71.
- Koo, B. W., Guhathakurt, S., & Botchwey, N. (2022). How are Neighborhood and Street-Level Walkability Factors Associated with Walking Behaviors? A Big Data Approach Using Street View Images. *ENVIRONMENT AND BEHAVIOR*, 54(1), 211–241. <https://doi.org/10.1177/00139165211014609>
- Koo, B. W., Guhathakurta, S., & Botchwey, N. (2022). Development and validation of automated microscale walkability audit method. *HEALTH & PLACE*, 73. <https://doi.org/10.1016/j.healthplace.2021.102733>
- Koohsari, M. J., Sugiyama, T., Mavoa, S., Villanueva, K., Badland, H., Giles-Corti, B., & Owen, N. (2016). Street network measures and adults' walking for transport: Application of space syntax. *Health and Place*, 38, 89–95. <https://doi.org/10.1016/j.healthplace.2015.12.009>

- Koskela, H., & Pain, R. (2000). Revisiting fear and place : women ' s fear of attack and the built environment. *Geoforum*, 31, 269–280.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in neural information processing systems*, 25(2), 1–1432. <https://doi.org/10.1201/9781420010749>
- Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the boruta package. *Journal of Statistical Software*, 36(11), 1–13. <https://doi.org/10.18637/jss.v036.i11>
- Kwon, K., & Akar, G. (2022). People with disabilities and use of public transit: The role of neighborhood walkability. *Journal of Transport Geography*, 100(February), 103319. <https://doi.org/10.1016/j.jtrangeo.2022.103319>
- Laine, J., & Kuvaja-ko, V. (2014). Cost-Effectiveness of Population-Level Physical Activity Interventions : A Systematic Review. *American Journal of Health Promotion*, 29(2), 71–81. <https://doi.org/10.4278/ajhp.131210-LIT-622>
- Larranaga, A. M., Arellana, J., Rizzi, L. I., Stramb, O., & Cybis, H. B. B. (2018). Using best – worst scaling to identify barriers to walkability : a study of Porto Alegre , Brazil. *Transportation*, 46, 2347–2379. <https://doi.org/10.1007/s11116-018-9944-x>
- Leão, A. L. F. (2017). Walkability: Objective and subjective measures of Brazilian towns' built environment attributes. *Associated Master Program in Architecture and Urbanism UEL/UEM*, 1–11.
- Leão, A. L. F. (2019). *Objective walkability measures for Brazilian towns: a methodological approach* [Masters Dissertation, Universidade Estadual de Londrina]. <https://doi.org/10.13140>
- Leão, A. L. F., & Kanashiro, M. (2021). Associações gênero-específicas da Caminhabilidade: Caminhada, uso do solo e características sociodemográficas. *Oculum Ensaios*, 19. <https://doi.org/10.24220/2318-0919v19e2022a5061>
- Leão, A. L. F., Lino, L., & Kanashiro, M. (2021). Women's walking by night: The influence of street connectivity and sociodemographic composition. *SIMPÓSIO NACIONAL DE GESTÃO E ENGENHARIA URBANA*, 441–447. Women's walking by night: The influence of street connectivity and sociodemographic composition.
- Leão, A. L. F., Queiroz, H., Barbon, S., & Kanashiro, M. (2020). Identificação de composições da paisagem urbana : uma abordagem de deep learning. *Revista de Morfologia Urbana*, 8(1), 2–14.
- Leão, A. L. F., & Urbano, M. R. (2020). *Street connectivity and walking: An empirical study in Londrina-PR*. June, 31–42. <https://doi.org/10.5433/1679-0375.2020v41n1p31>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. <https://doi.org/10.1038/nature14539>
- Lee, C., & Moudon, A. V. (2006). Correlates of walking for transportation or recreation purposes. *Journal of Physical Activity & Health*, 3, 77–98. <https://doi.org/10.1016/j.trd.2006.02.003>
- Lee, S. M. (2010). *The Association between Neighborhood Environment and Neighborhood Satisfaction: Moderating Effects of Demographics (Thesis)*. San Diego State University.
- Lee, S., & Talen, E. (2014). Measuring Walkability : A Note on Auditing Methods. *Journal of Urban Design*, 19(3), 368–388. <https://doi.org/10.1080/13574809.2014.890040>
- Leslie, E., Cerin, E., DuToit, L., Owen, N., & Bauman, A. (2007). Objectively Assessing 'Walkability' of Local Communities: Using GIS to Identify the Relevant Environmental Attributes. Em A. Lai, PC and Mak (Org.), In: *Lai P.C., Mak A.S.H. (eds) GIS for Health and the Environment. Lecture Notes in Geoinformation and Cartography*. (1º ed). Springer.
- Leslie, E., Coffee, N., Frank, L., Owen, N., Bauman, A., & Hugo, G. (2007). Walkability of local communities: Using geographic information systems to objectively assess relevant

- environmental attributes. *Health and Place*, 13(1), 111–122.  
<https://doi.org/10.1016/j.healthplace.2005.11.001>
- Leyden, K. M. (2003). Social Capital and the Built Environment: The Importance of Walkable Neighborhoods. *American Journal of Public Health*, 93(9), 1546–1551.  
<https://doi.org/10.2105/AJPH.93.9.1546>
- Leyland AH, & Groenewegen PP. (2020). *Multilevel Modelling for Public Health and Health Services Research: Health in Context* (PMID: 33347097.). Cham (CH): Springer; .
- Li, X., Zhang, C., Li, W., Ricard, R., & Meng, Q. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening*, 14(3), 675–685. <https://doi.org/10.1016/j.ufug.2015.06.006>
- Li, Y., Peng, L., Wu, C., & Zhang, J. (2022). Street View Imagery (SVI) in the Built Environment: A Theoretical and Systematic Review. *BUILDINGS*, 12(8).  
<https://doi.org/10.3390/buildings12081167>
- Liang, J., Gong, J., Sun, J., Zhou, J., Li, W., & Li, Y. (2017). Automatic Sky View Factor Estimation from Street View Photographs — A Big Data Approach. *remote sensing*, 9(411), 1–17. <https://doi.org/10.3390/rs9050411>
- Liu, L., Silva, E. A., Wu, C., & Wang, H. (2017a). A machine learning-based method for the large-scale evaluation of the qualities of the urban environment. *Computers, Environment and Urban Systems*, 65, 113–125.  
<https://doi.org/10.1016/j.compenvurbsys.2017.06.003>
- Liu, L., Silva, E. A., Wu, C., & Wang, H. (2017b). A machine learning-based method for the large-scale evaluation of the qualities of the urban environment. *Computers, Environment and Urban Systems*, 65, 113–125.  
<https://doi.org/10.1016/j.compenvurbsys.2017.06.003>
- Long, Y., & Liu, L. (2017a). *How green are the streets? An analysis for central areas of Chinese cities using Tencent Street View*. 1–18.  
<https://doi.org/10.1371/journal.pone.0171110>
- Long, Y., & Liu, L. (2017b). *How green are the streets? An analysis for central areas of Chinese cities using Tencent Street View*. 1–18.  
<https://doi.org/10.1371/journal.pone.0171110>
- Lotfi, S., & Koochsari, M. J. (2011). Neighborhood Walkability in a City within a Developing Country. *Journal of Urban Planning and Development*, 137(4), pp 402-408.  
[http://dx.doi.org/10.1061/\(ASCE\)UP.1943-5444.0000085](http://dx.doi.org/10.1061/(ASCE)UP.1943-5444.0000085)
- Lovasi, G. S., Moudon, A. V., Pearson, A. L., Hurvitz, P. M., Larson, E. B., Siscovick, D. S., Berke, E. M., Lumley, T., & Psaty, B. M. (2008). Using built environment characteristics to predict walking for exercise. *INTERNATIONAL JOURNAL OF HEALTH GEOGRAPHICS*, 7. <https://doi.org/10.1186/1476-072X-7-10>
- Lu, Y. (2018a). The Association of Urban Greenness and Walking Behavior: Using Google Street View and Deep Learning Techniques to Estimate Residents' Exposure to Urban Greenness. *International Journal of Environmental Research and Public Health Article*, 15(8), 1–15. <https://doi.org/10.3390/ijerph15081576>
- Lu, Y. (2018b). The association of urban greenness and walking behavior: Using google street view and deep learning techniques to estimate residents' exposure to urban greenness. *International Journal of Environmental Research and Public Health*, 15(8).  
<https://doi.org/10.3390/ijerph15081576>
- Lu, Y. (2018c). The association of urban greenness and walking behavior: Using google street view and deep learning techniques to estimate residents' exposure to urban greenness. *International Journal of Environmental Research and Public Health*, 15(8).  
<https://doi.org/10.3390/ijerph15081576>

- Luan, H., Ramsay, D., & Fuller, D. (2019). Household income , active travel , and their interacting impact on body mass index in a sample of urban Canadians : a Bayesian spatial analysis. *International Journal of Health Geographics*, 18(4), 1–10. <https://doi.org/10.1186/s12942-019-0168-x>
- Lucchesi, S. T., E Silva, J. de A., Larranaga, A. M., Zechin, D., & Cybis, H. B. B. (2023). Machine Learning and Image Recognition Technologies to Identify Built Environment Barriers and Incentives to Walk. *Em Transportation Research Record* (Vol. 2677, Número 1, p. 14–24). SAGE Publications Ltd. <https://doi.org/10.1177/03611981221097965>
- Lukka, K. (2003). The constructive research approach. *Case study research in logistics. Publications of the Turku School of Economics and Business Administration, Series B1*, 83–101.
- Macintyre, S., Ellaway, A., & Cummins, S. (2002). Place effects on health: how can we conceptualise, operationalise and measure them? *SOCIAL SCIENCE & MEDICINE*, 55(1), 125–139. [https://doi.org/10.1016/S0277-9536\(01\)00214-3](https://doi.org/10.1016/S0277-9536(01)00214-3)
- Mackenbach, J. D., Lakerveld, J., van Lenthe, F. J., Bárdos, H., Glonti, K., Compernelle, S., De Bourdeaudhuij, I., Oppert, J. .-M., Roda, C., Rutter, H., Brug, J., & Nijpels, G. (2016). Exploring why residents of socioeconomically deprived neighbourhoods have less favourable perceptions of their neighbourhood environment than residents of wealthy neighbourhoods. *Obesity Reviews*, 17, 42–52. <https://doi.org/10.1111/obr.12375>
- Maghelal, P. K., & Capp, C. J. (2017). Walkability : A Review of Existing Pedestrian Indices. *Journal of the Urban and Regional Information Systems Association, January 2011*.
- March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(1), 251–266.
- Marijana, Ć. (2020). CNN Classification of the Cultural Heritage Images. *19th International Symposium INFOTEH-JAHORINA, 18-20 March 2020, March*, 18–20.
- Martin, D. (1998). *Developing the Automated Zoning Procedure to Reconcile Incompatible Zoning Systems. 1995*.
- McCormack, G. R., Cerin, E., Leslie, E., Du Toit, L., & Owen, N. (2008). Objective versus perceived walking distances to destinations - Correspondence and predictive validity. *ENVIRONMENT AND BEHAVIOR*, 40(3), 401–425. <https://doi.org/10.1177/0013916507300560>
- Middel, A., Lukasczyk, J., Zakrzewski, S., Arnold, M., & Maciejewski, R. (2019). Urban form and composition of street canyons : A human-centric big data and deep learning approach. *Landscape and Urban Planning*, 183, 122–132. <https://doi.org/10.1016/j.landurbplan.2018.12.001>
- Millstein, R. A., Cain, K. L., Sallis, J. F., Conway, T. L., Geremia, C., Frank, L. D., Chapman, J., Dyck, D. Van, Dipzinski, L. R., Kerr, J., Glanz, K., & Saelens, B. E. (2013). *Development , scoring , and reliability of the Microscale Audit of Pedestrian Streetscapes ( MAPS )*.
- Moniruzzaman, Md., & Paez, A. (2012). A model-based approach to select case sites for walkability audits. *Health & Place*, 18, 1323–1334. <https://doi.org/10.1016/j.healthplace.2012.09.013>
- Moosavi, V. (2017). Urban morphology meets deep learning : Exploring urban forms in one million cities , town and villages across the planet. *arXiv preprint, arXiv:1709*, 1–10.
- Moreno, C., Allam, Z., Chabaud, D., Gall, C., & Pralong, F. (2021). Introducing the “15-minute city”: Sustainability, resilience and place identity in future post-pandemic cities. *Smart Cities*, 4(1), 93–111. <https://doi.org/10.3390/smartcities4010006>
- MOTOMURA, M. C. N. (2017). Walkable spaces: correlation between built environment, social capital and self-rated health in Cambé – Paraná – Brazil. *Associated Master*

- Program in Architecture and Urbanism from State University of Londrina and State University of Maringá., Dissertação de Mestrado., 170.*
- Moudon, A. V., & Lee, C. (2003). Walking and Biking: An Evaluation of Environmental Audit Instruments. *American Journal of Health Promotion, 18*(No. 1), 21–37. <https://doi.org/10.4278/0890-1171-18.1.21>
- Murray, C., Lopez, A. D., & Cambridge, M. A. (2014). Transport for Health: The global burden of disease from motorized road transport. Em *The world bank; Institute for health metrics and Evaluation*. (Vol. 1, Número 1).
- Nguyen, Q. C., Huang, Y., Kumar, A., Duan, H., & Keralis, J. M. (2020). Using 164 Million Google Street View Images to Derive Built Environment Predictors of COVID-19 Cases. *July*, 1–13.
- Novack, T., Vorbeck, L., Lorei, H., & Zipf, A. (2020). Towards Detecting Building Facades with Graffiti Artwork Based on Street View Images. <https://doi.org/10.3390/ijgi9020098>
- Oakes, J. M., Forsyth, A., & Schmitz, K. H. (2007). The effects of neighborhood density and street connectivity on walking behavior: The Twin Cities walking study. *Epidemiologic Perspectives and Innovations, 4*. <https://doi.org/10.1186/1742-5573-4-16>
- Obeso, A. M., & Benois-pineau, J. (2019). Architectural style classification of Mexican historical buildings using deep convolutional neural networks and sparse features. *Journal of Electronic Imaging, 26*(1). <https://doi.org/10.1117/1.JEI.26.1.011016>
- Olsen, J. R., Mitchell, R., Mutrie, N., Foley, L., & Ogilvie, D. (2017). Population levels of , and inequalities in , active travel : A national , cross- sectional study of adults in Scotland. *Preventive Medicine Reports, 8*(September), 129–134. <https://doi.org/10.1016/j.pmedr.2017.09.008>
- Openshaw, S. (1984). *MODIFIABLE AREAL UNIT PROBLEM* (C. and T. in M. G.-CATMOG, Org.; 1º ed). Concepts and techniques in Modern Geography No.38. Geobooks, Norwich.
- Ord, J. K., & Getis, A. (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis, 27*(4), 286–306. <https://doi.org/10.1111/j.1538-4632.1995.tb00912.x>
- Orellana, D., Hermida, C., & Osorio, P. (2016). A multidisciplinary analytical framework for studying active mobility patterns. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 41*(July), 527–534. <https://doi.org/10.5194/isprsarchives-XLI-B2-527-2016>
- Owen, N., Cerin, E., Leslie, E., duToit, L., Coffee, N., Frank, L. D., Bauman, A. E., Hugo, G., Saelens, B. E., & Sallis, J. F. (2007). Neighborhood Walkability and the Walking Behavior of Australian Adults. *American Journal of Preventive Medicine, 33*(5), 387–395. <https://doi.org/10.1016/j.amepre.2007.07.025>
- Pain, R. (2001). Gender, Race, Age and Fear in the City. *Urban Studies, 38*(5–6), 899–913. <https://doi.org/10.1080/00420980120046590>
- Pain, R. H. (1997). Social Geographies of Women’s Fear of Crime. *The Royal Geographical Society (with the Institute of British Geographers), 22*(2), 231–244.
- Pani-Harremán, K. E., Bours, G. J. J. W., Zander, I., Kempen, G. I. J. M., & Van Duren, J. M. A. (2021). Definitions, key themes and aspects of “ageing in place”: A scoping review. Em *Ageing and Society* (Vol. 41, Número 9, p. 2026–2059). Cambridge University Press. <https://doi.org/10.1017/S0144686X20000094>
- Park, S., Choi, K., & Lee, J. S. (2015). To Walk or Not to Walk: Testing the Effect of Path Walkability on Transit Users’ Access Mode Choices to the Station. *International Journal of Sustainable Transportation, 9*(8), 529–541. <https://doi.org/10.1080/15568318.2013.825036>

- Park, S., Deakin, E., & Lee, J. S. (2014). Perception-Based Walkability Index to Test Impact of Microlevel Walkability on Sustainable Mode Choice Decisions. *Transportation Research Record*, 2464(1), 126–134. <https://doi.org/10.3141/2464-16>
- Pereira, R. H. M., & Karner, A. (2021). Transportation Equity. *International Encyclopedia of Transportation*, 1, 271–277. <https://doi.org/10.1016/b978-0-08-102671-7.10053-3>
- Pereira, R. H. M., Schwanen, T., & Banister, D. (2017). Distributive justice and equity in transportation. *Transport Reviews*, 37(2), 170–191. <https://doi.org/10.1080/01441647.2016.1257660>
- Pikora, T. (2000). *Survey of the Physical Environment in Local Neighbourhoods - SPACES Instrument*.
- Pikora, T. J., Bull, F. C. L., Jamrozik, K., Knuiaman, M., Giles-Corti, B., & Donovan, R. J. (2002). Developing a reliable audit instrument to measure the physical environment for physical activity. *American Journal of Preventive Medicine*, 23(3), 187–194. [https://doi.org/10.1016/S0749-3797\(02\)00498-1](https://doi.org/10.1016/S0749-3797(02)00498-1)
- Pliakas, T., Hawkesworth, S., Silverwood, R. J., Nanchahal, K., Grundy, C., Armstrong, B., Casas, J. P., Morris, R. W., Wilkinson, P., & Lock, K. (2017). Optimising measurement of health-related characteristics of the built environment: Comparing data collected by foot-based street audits, virtual street audits and routine secondary data sources. *Health & Place*, 43(October 2016), 75–84. <https://doi.org/10.1016/j.healthplace.2016.10.001>
- Poggio, T., Mhaskar, H., Rosasco, L., Miranda, B., & Liao, Q. (2017). Why and When Can Deep – but Not Shallow – Networks Avoid the Curse of Dimensionality: a Review. *International Journal of Automation and Computing*, 14(6), 503–519. <https://doi.org/10.1007/s11633-017-1054-2>
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <http://www.r-project.org/>
- Rachele, J. N., Learnihan, V., Badland, H. M., Mavoa, S., Turrell, G., & Giles-Corti, B. (2018). Are measures derived from land use and transport policies associated with walking for transport? *Journal of Physical Activity and Health*, 15(1), 1–9. <https://doi.org/10.1123/jpah.2016-0693>
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2019). *Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer*. <http://arxiv.org/abs/1910.10683>
- Rego, R. L., & Meneguetti, K. S. (2006). A forma urbana das cidades de médio porte e dos patrimônios fundados pela Companhia Melhoramentos Norte do Paraná. *Acta Scientiarum. Technology*, 28(1). <https://doi.org/10.4025/actascitechnol.v28i1.1292>
- Reis, R. S., Akira, A., & Hino, F. (2013). Walkability and Physical Activity: Findings from Curitiba, Brazil. *American Journal of Preventive Medicine*, 45(3), 269–275.
- Rundle, A. G., Bader, M. D. M., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using Google Street View to Audit Neighborhood Environments. *AMEPRE*, 40(1), 94–100. <https://doi.org/10.1016/j.amepre.2010.09.034>
- Rzotkiewicz, A., Pearson, A. L., Dougherty, B. V, Shortridge, A., & Wilson, N. (2018a). Systematic review of the use of Google Street View in health research : Major themes , strengths , weaknesses and possibilities for future research. *Health & Place*, 52(September 2017), 240–246. <https://doi.org/10.1016/j.healthplace.2018.07.001>
- Rzotkiewicz, A., Pearson, A. L., Dougherty, B. V, Shortridge, A., & Wilson, N. (2018b). Systematic review of the use of Google Street View in health research : Major themes , strengths , weaknesses and possibilities for future research ☆. *Health & Place*, 52(July), 240–246. <https://doi.org/10.1016/j.healthplace.2018.07.001>
- Sabel, C. E., Kihal, W., Bard, D., & Weber, C. (2013). Creation of synthetic homogeneous neighbourhoods using zone design algorithms to explore relationships between asthma

- and deprivation in Strasbourg, France. *Social Science and Medicine*, 91, 110–121.  
<https://doi.org/10.1016/j.socscimed.2012.11.018>
- Saelens, B. E., & Sallis, J. F. (2002). Neighborhood Environment Walkability Scale (NEWS). *American Journal of Public Health*, 93(9), 78–81.  
<https://doi.org/10.1016/j.yjmed.2009.07.011>
- SAELEN, B. E. ;, & SALLIS, J. F. (2002). Neighborhood Environment Walkability Scale (NEWS). *American Journal of Public Health*, 93(9), 78–81.  
<https://doi.org/10.1080/15575339609489851>
- Saelens, B. E., Sallis, J. F., Black, J. B., & Chen, D. (2003). Neighborhood-based differences in physical activity: An environment scale evaluation. *AMERICAN JOURNAL OF PUBLIC HEALTH*, 93(9), 1552–1558. <https://doi.org/10.2105/AJPH.93.9.1552>
- Saelens, B. E., Sallis, J. F., & Frank, L. D. (2003). Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures. *Annals of Behavioral Medicine*, 25(2), 80–91. [https://doi.org/10.1207/S15324796ABM2502\\_03](https://doi.org/10.1207/S15324796ABM2502_03)
- Saelens, B., & Handy, S. L. (2008). Built environment correlates of walking: a review. *Med Sci Sports Exerc*, 40. <https://doi.org/10.1249/MSS.0b013e31817c67a4>
- Sallis, J. F. (2016). *Microscale Audit of Pedestrian Streetscapes (MAPS) Global Survey*. University of California. [http://sallis.ucsd.edu/measure\\_maps.html](http://sallis.ucsd.edu/measure_maps.html)
- Sallis, J. F., Bull, F., Burdett, R., Frank, L. D., Griffiths, P., Giles-Corti, B., & Stevenson, M. (2016). Use of science to guide city planning policy and practice: how to achieve healthy and sustainable future cities. *The Lancet*, 388(10062), 2936–2947.  
[https://doi.org/10.1016/S0140-6736\(16\)30068-X](https://doi.org/10.1016/S0140-6736(16)30068-X)
- Sallis, J. F., Owen, N., & Fisher, E. B. (2008). Ecological models of health behavior. In *Health behavior and health education: Theory, research, and practice* (p. 465–485). Jossey-Bass.
- Sallis, J. F., Slymen, D. J., Conway, T. L., Frank, L. D., Saelens, B. E., Cain, K., & Chapman, J. E. (2011). Income disparities in perceived neighborhood built and social environment attributes. *Health and Place*, 17(6), 1274–1283.  
<https://doi.org/10.1016/j.healthplace.2011.02.006>
- Salvo, D., Jáuregui, A., Adlakha, D., Sarmiento, O. L., & Reis, R. S. (2023). When moving is the only option: the role of necessity versus choice for understanding and promoting physical activity in low-and middle-income countries. *Annu Rev Public Health* ., 44, 151–169.
- Sattler, M., Köberl, G., Färber, T., Katharina Traußnig, J. J., Dietz, P., & Poppel, M. van. (2018). Healthy on the way (HOTway): Perceived stress of different modes of commuting. *7th International Society for Physical Activity and Health Congress*, 15(October), 1–251.
- Schwanen, T., Kwan, M., & Ren, F. (2008). How fixed is fixed? Gendered rigidity of space – time constraints and geographies of everyday activities. *Geoforum*, 39(6), 2109–2121.  
<https://doi.org/10.1016/j.geoforum.2008.09.002>
- Seiferling, I., Naik, N., Ratti, C., & Proulx, R. (2017). Green streets – Quantifying and mapping urban trees with street-level imagery and computer vision. *Landscape and Urban Planning*, 165(May), 93–101. <https://doi.org/10.1016/j.landurbplan.2017.05.010>
- Shalunts, G., Haxhimusa, Y., & Sablatnig, R. ([s.d.]). *Architectural Style Classification of Building Facade Windows*. 1–10.
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Shashank, A., & Schuurman, N. (2019). Unpacking walkability indices and their inherent assumptions. *Health & Place*, 55(November 2018), 145–154.  
<https://doi.org/10.1016/j.healthplace.2018.12.005>

- Shelhamer, E., Long, J., & Darrell, T. (2017). Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4), 640–651. <https://doi.org/10.1109/TPAMI.2016.2572683>
- Shen, Q., Member, S., Zeng, W., Ye, Y., Stefan, M., Schubiger, S., Burkhard, R., & Qu, H. (2018). StreetVizor : Visual Exploration of Human-Scale Urban Forms Based on Street Views. *IEEE Transactions On Visualization And Computer Graphics*, 24(1), 1004–1013. <https://doi.org/10.1109/TVCG.2017.2744159>
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–14.
- Steinmetz-wood, M., Velauthapillai, K., Brien, G. O., & Ross, N. A. (2019). *Assessing the micro-scale environment using Google Street View : the Virtual Systematic Tool for Evaluating Pedestrian Streetscapes ( Virtual-STEPS )*. 3, 1–11.
- Stockton, J. C., Duke-williams, O., Stamatakis, E., Mindell, J. S., Brunner, E. J., & Shelton, N. J. (2016). *Development of a novel walkability index for London , United Kingdom : cross- sectional application to the Whitehall II Study*. 1–13. <https://doi.org/10.1186/s12889-016-3012-2>
- Sugiyama, T., Neuhaus, M., Cole, R., Giles-Corti, B., & Owen, N. (2012). Destination and Route Attributes Associated with Adults' Walking: A Review. *MEDICINE AND SCIENCE IN SPORTS AND EXERCISE*, 44(7), 1275–1286. <https://doi.org/10.1249/MSS.0b013e318247d286>
- Sugiyama, T., Rachele, J. N., Gunn, L. D., Burton, N. W., Brown, W. J., & Turrell, G. (2019). Land use proportion and walking :Application of isometric substitution analysis. *Health & Place*, 57, 352–357. <https://doi.org/10.1016/j.healthplace.2018.12.004>
- Talen, E. (2002). Pedestrian access as a measure of urban quality. *Planning Practice and Research*, 17(3), 257–278. <https://doi.org/10.1080/026974502200005634>
- Tan, Y., Tang, P., Zhou, Y., Luo, W., Kang, Y., & Li, G. (2017). Neurocomputing Photograph aesthetical evaluation and classification with deep convolutional neural networks. *Neurocomputing*, 228(August 2016), 165–175. <https://doi.org/10.1016/j.neucom.2016.08.098>
- Tilt, J. H., Unfried, T. M., & Roca, B. (2007). Using objective and subjective measures of neighborhood greenness and accessible destinations for understanding walking trips and BMI in Seattle, Washington. *American Journal of Health Promotion*, 21(4 SUPPL.), 371–379.
- Tobin, M., Hajna, S., Orychock, K., Ross, N., DeVries, M., Villeneuve, P. J., Frank, L. D., McCormack, G. R., Wasfi, R., Steinmetz-Wood, M., Gilliland, J., Booth, G. L., Winters, M., Kestens, Y., Manaugh, K., Rainham, D., Gauvin, L., Widener, M. J., Muhajarine, N., ... Fuller, D. (2022). Rethinking walkability and developing a conceptual definition of active living environments to guide research and practice. *BMC Public Health*, 22(1), 1–7. <https://doi.org/10.1186/s12889-022-12747-3>
- Vaishnavi, V. K., & Jr., W. K. (2008). *Design Science Research Methods and Patterns Innovating Information and Communication Technology*. Auerbach Publications Taylor & Francis Group.
- Van Cauwenberg, J., Van Holle, V., De Bourdeaudhuij, I., Van Dyck, D., & Deforche, B. (2016). Neighborhood walkability and health outcomes among older adults: The mediating role of physical activity. *Health & place*, 37, 16–25. <https://doi.org/10.1016/j.healthplace.2015.11.003>
- Van Cauwenberg, J., Van Holle, V., Simons, D., Deridder, R., Clarys, P., Goubert, L., Nasar, J., Salmon, J., De Bourdeaudhuij, I., & Deforche, B. (2012). Environmental Factors Influencing Older Adults' Walking for Transportation: A Study Using Walk-along

- Interviews. *International Journal of Behavioral Nutrition and Physical Activity*, 9(1), 11p.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*. <http://arxiv.org/abs/1706.03762>
- Wang, H., & Yang, Y. (2019a). *Neighbourhood walkability : A review and bibliometric analysis*. 93(April), 43–61. <https://doi.org/10.1016/j.cities.2019.04.015>
- Wang, H., & Yang, Y. (2019b). *Neighbourhood walkability : A review and bibliometric analysis*. *Cities*, 93(September 2018), 43–61. <https://doi.org/10.1016/j.cities.2019.04.015>
- Wang, W., Xie, E., Li, X., Fan, D.-P., Song, K., Liang, D., Lu, T., Luo, P., & Shao, L. (2021). Pyramid Vision Transformer: A Versatile Backbone for Dense Prediction without Convolutions. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 548–558. <https://doi.org/10.1109/ICCV48922.2021.00061>
- Watts, N., Adger, W. N., Agnolucci, P., & Blackstock, J. (2015). Health and climate change: Policy responses to protect public health. *Environnement, Risques et Sante*, 14(6), 466–468. [https://doi.org/10.1016/S0140-6736\(15\)60854-6](https://doi.org/10.1016/S0140-6736(15)60854-6)
- Wesemael, V. (2020). How does walkability change behavior ? A Comparison between Different Age Groups. *International Journal of Environmental Research and Public Health*, 17(540), 1–14. <https://doi.org/10.3390/ijerph17020540>
- Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). *SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers*. <http://arxiv.org/abs/2105.15203>
- Xu, Z., Tao, D., Zhang, Y., Wu, J., & Tsoi, A. C. ([s.d.]). *Architectural Style Classification using Multinomial Latent Logistic Regression*. 1–16.
- Yang, Y., Diez Roux, A. V., Auchincloss, A. H., Rodriguez, D. A., & Brown, D. G. (2012). Exploring walking differences by socioeconomic status using a spatial agent-based model. *Health and Place*, 18(1), 96–99. <https://doi.org/10.1016/j.healthplace.2011.08.010>
- Yin, L., & Wang, Z. (2016a). Measuring visual enclosure for street walkability : Using machine learning algorithms and Google Street View imagery. *Applied Geography*, 76, 147–153. <https://doi.org/10.1016/j.apgeog.2016.09.024>
- Yin, L., & Wang, Z. (2016b). Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. *Applied Geography*, 76, 147–153. <https://doi.org/10.1016/j.apgeog.2016.09.024>
- Yin, L., & Wang, Z. (2016c). Measuring visual enclosure for street walkability : Using machine learning algorithms and Google Street View imagery. *Applied Geography*, 76, 147–153. <https://doi.org/10.1016/j.apgeog.2016.09.024>
- Yin, R. K. (2001). *Estudo de caso: Planejamento e Métodos* (C. DAMACENA, Org.; 2º). BOOKMAN COMPANHIA EDITORA.
- Yue, X., Antonietti, A., Alirezai, M., Tasdizen, T., Li, D., Nguyen, L., Mane, H., Sun, A., Hu, M., Whitaker, R. T., & Nguyen, Q. C. (2022). Using Convolutional Neural Networks to Derive Neighborhood Built Environments from Google Street View Images and Examine Their Associations with Health Outcomes. *International Journal of Environmental Research and Public Health*, 19(19). <https://doi.org/10.3390/ijerph191912095>
- Zampieri, F. L. L. (2012). O fenômeno social do movimento de pedestres em centros urbanos. *Propur, Doutorado*, 928.
- Zeng, L., Lu, J., Li, W., & Li, Y. (2018). A fast approach for large-scale Sky View Factor estimation using street view images. *Building and Environment*, 135(December 2017), 74–84. <https://doi.org/10.1016/j.buildenv.2018.03.009>

- Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning*, 180(October 2017), 148–160. <https://doi.org/10.1016/j.landurbplan.2018.08.020>
- Zhang, Y., & Dong, R. (2018). Impacts of Street-Visible Greenery on Housing Prices : Evidence from a Hedonic Price Model and a Massive Street View Image Dataset in Beijing. *International Journal of Geo-Information*, 7(104), 1–19. <https://doi.org/10.3390/ijgi7030104>
- Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid Scene Parsing Network Hengshuang. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.
- Zheng, S., Lu, J., Zhao, H., Zhu, X., Luo, Z., Wang, Y., Fu, Y., Feng, J., Xiang, T., Torr, P. S., & Zhang, L. (2021). Rethinking Semantic Segmentation from a Sequence-to-Sequence Perspective with Transformers. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 6877–6886. <https://doi.org/10.1109/CVPR46437.2021.00681>
- Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., & Torralba, A. (2018). Places: A 10 Million Image Database for Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(6), 1452–1464. <https://doi.org/10.1109/TPAMI.2017.2723009>
- Zhou, B., Zhao, H., Puig, X., Xiao, T., Fidler, S., Barriuso, A., & Torralba, A. (2019). Semantic Understanding of Scenes Through the ADE20K Dataset. *International Journal of Computer Vision*, 127(3), 302–321. <https://doi.org/10.1007/s11263-018-1140-0>
- Zhou, H., He, S., Cai, Y., Wang, M., & Su, S. (2019). Social inequalities in neighborhood visual walkability : Using street view imagery and deep learning technologies to facilitate healthy city planning. *Sustainable Cities and Society*, 50(129), 101605. <https://doi.org/10.1016/j.scs.2019.101605>