A Computer Vision–Enriched Discrete Choice Model for Pedestrian Route Choice Preferences in the Netherlands

Integrating Street-Level Visual Characteristics into Pedestrian Route Choice Modelling across Trip-Purpose Contexts

MSc Engineering and Policy Analysis Thesis

P.J.M. Kastelein

Delft University of Technology

A Computer Vision–Enriched Discrete Choice Model for Pedestrian Route Choice Preferences in the Netherlands

Integrating Street-Level Visual Characteristics into Pedestrian Route Choice Modelling across Trip-Purpose Contexts

by

P.J.M. Kastelein

to obtain the degree of Master of Science
in Engineering and Policy Analysis,
at the Delft University of Technology,
to be defended publicly on Monday July 6, 2026 at 13:30

Student number:	4666135	
Project duration:	February 2, 2026 – July 6, 2026	
Thesis committee:	Dr. Sander van Cranenburgh	TU Delft, Chair
	Dr. Natalie van der Wal	TU Delft, Supervisor
	Dr. Ali Nadi	TNO SUMS, Supervisor

Preface

After almost nine years of studying, this thesis marks the end of my student years. Looking back, my academic journey has not always been straightforward, but I am grateful for everything I have learned and the experiences I have had. My years in the Computer Science bachelor brought me a lot of knowledge, many good memories with my friends and roommates, and some fun years of rowing and coaching at Proteus. The past two years in Engineering and Policy Analysis introduced me to a field that genuinely interests me and to some amazing new people. At the same time, losing my father in my first year made parts of this journey more difficult. So reaching this point is something I am very proud of, and I think he would have been proud too.

This thesis allowed me to bring together both sides of my background. It introduced me to the field of choice modelling, challenged me to go outside my comfort zone by designing and conducting a stated preference survey, and gave me the opportunity to explore how computer science methods can be applied to questions about human behaviour and urban environments. Modelling and simulating behaviour is something I have become very interested in, and I hope to continue exploring it in the future. Although these past months of working on my thesis were sometimes stressful, I also really enjoyed learning so many new things. I am happy with the final result and I think it reflects what I have learned throughout my studies.

First, I would like to thank Sander for having confidence in me from the start, even though I had no prior experience with choice modelling. Thank you for making it possible for me to work with TNO and for giving me the opportunity to take on this challenge. I really appreciated our technical discussions, in which you were always able to explain difficult concepts very clearly. These discussions helped me understand the theory and modelling choices much better.

I would also like to thank Natalie for the many behavioural insights you shared with me, and for your tips on the survey design. You helped me to really improve the survey and reach many more respondents. Most of all, thank you for your enthusiasm and your good advice.

I also want to thank Ali, my supervisor at TNO, for making it possible for me to do this project and for guiding me throughout the process. During our weekly meetings, I always felt comfortable asking questions, and you helped me organise my sometimes confused thoughts. You also reminded me that it is okay to take things a bit slower sometimes. Your kind words really gave me confidence in my work.

I have really enjoyed my time at TNO, and I would like to thank everyone there for making it such a nice and inspiring place to work. I especially enjoyed the lunch walks, Kennislunches, and discussions with the other thesis interns. These moments made the thesis process a lot more fun and often gave me useful new ideas.

Finally, I would like to thank my boyfriend for his support during this period. His positive attitude made the whole process a lot easier, and I really appreciate the energy and encouragement he gave me when I needed it the most. I also want to thank my friends and family for their support over the past months, and for making my student years at TU Delft so much fun. I am grateful for everything I have learned along the way, and I look forward to what comes next.

Thank you all,
Pien Kastelein
Delft, July 2026

Summary

Walking plays an important role in Dutch urban mobility, both for short everyday trips and as part of longer journeys to and from public transport stops or stations. However, pedestrian route choice is often simplified in transport models, for example by assuming that pedestrians choose the shortest or fastest route. This simplification matters for urban policy, because walking is increasingly linked to sustainable mobility, public health, and more efficient use of public space. If pedestrian route choice is mainly represented through distance or travel time, models may overlook visual street-level qualities that are relevant for evaluating walking environments and pedestrian-oriented policies. These qualities, such as greenery, visible traffic, pedestrian space, and the presence of other people, may influence whether a route is perceived as attractive, comfortable, or safe. The importance of these qualities may also differ by trip purpose. For example, someone walking to public transport may value travel time more strongly than someone walking in their free time. This thesis therefore investigates to what extent visual and non-visual route attributes influence pedestrian route choice preferences in the Netherlands, and how these preferences differ across trip-purpose contexts. The research was conducted in collaboration with the Sustainable Urban Mobility and Space (SUMS) department of TNO and focuses on the use of street-level visual information in pedestrian route choice modelling.

To study these preferences, a Stated Preference (SP) experiment was designed in which respondents repeatedly chose between two walking route alternatives that differed in travel time and the visual appearance of the street, shown through a street-level image. The choice tasks were presented in four trip-purpose contexts: walking to public transport, walking to work or school, walking home, and walking in free time. The collected data were used to estimate three pedestrian route choice model specifications for each trip purpose. As shown in Figure 1, these models differ in how visual information is included and in the trade-off between interpretability and flexibility. The baseline Multinomial Logit (MNL) only includes travel time, the pixel-share MNL uses predefined visual attributes extracted from the images, and the Computer Vision-Enriched Discrete Choice Model (CV-DCM) uses a deep neural network to learn visual information directly from the full street-level image. This made it possible to compare different levels of visual representation.

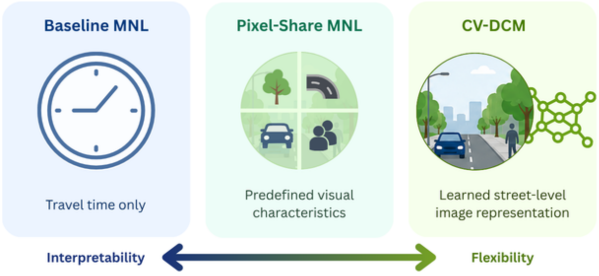


Figure 1: Overview of the three model specifications and the trade-off between interpretability and flexibility.

The results show that travel time remains an important factor in pedestrian route choice preferences, especially for more goal-oriented trips. However, travel time alone does not fully explain the observed choices. In the collected SP dataset, models that include visual information generally perform better than the baseline model, with the CV-DCM showing the strongest predictive performance in most contexts. In the pixel-share MNL, visible greenery was associated with the clearest and most consistent positive effect on route preferences. The qualitative validation of the CV-DCM also suggested that images with higher predicted utility were often greener and more attractive. At the same time, the importance of visual information differs across trip purposes. As summarised in Figure 2, the relative importance of travel time and visual information changes across the four walking contexts.

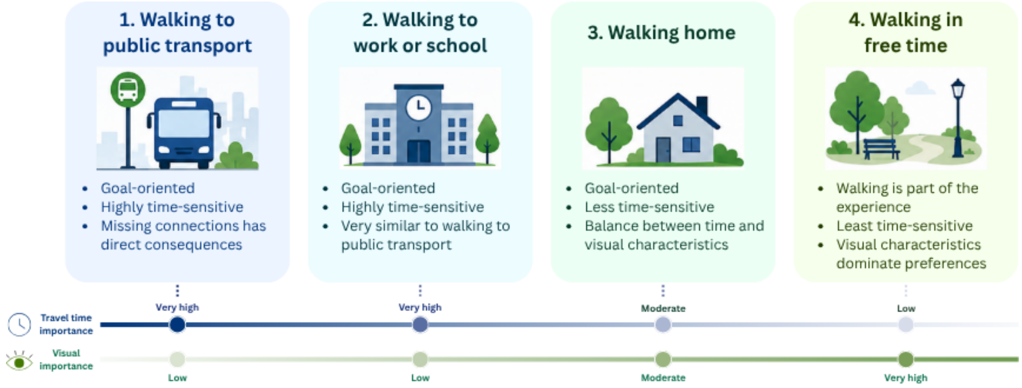


Figure 2: Summary of the four trip-purpose contexts and the relative importance of travel time and visual information.

The trip-purpose-specific CV-DCM models were also used in an exploratory analysis of a pedestrian network in Amsterdam-Zuid. Street-level image utilities were linked to street segments and combined with travel time utility to compare the shortest route with the route selected by each model. In this exploratory application, the model-selected routes were not always the shortest routes and also differed between trip-purpose models. This illustrates how visual route preferences can be mapped across a pedestrian network and used to explore how street-level visual quality may influence route attractiveness in different walking contexts.

The findings of this thesis have several implications for pedestrian route choice modelling and urban policy. Models based only on travel time may be too limited when the aim is to explain pedestrian route preferences or evaluate walking environments. However, this does not mean that every pedestrian model needs to include full street-level images or a complex CV-DCM. Instead, the level of model complexity should match the purpose of the analysis. A simpler shortest-path model may be sufficient for a basic approximation of pedestrian movement, while visual street-level information becomes more relevant when the aim is to evaluate walkability, route attractiveness, or street-level interventions. The results also show that trip purpose matters. Walking to public transport, walking to work or school, walking home, and walking in free time should not always be represented as one general walking context, because the relative importance of travel time and visual street-level characteristics differs between these trip purposes.

At the same time, the results should be interpreted within the limitations of the research design. The study uses an SP experiment, which made it possible to measure controlled trade-offs between travel time and visual street-level information. However, stated choices do not fully represent actual walking behaviour. Respondents did not physically experience the routes or the additional walking time, and the survey sample was not representative of the full Dutch population. The route alternatives were also simplified, because each route alternative was represented only by travel time and one street-level image, while real walking routes consist of multiple street segments and changing visual conditions. In addition, the CV-DCM generally improves predictive performance, but its learned image representation is less directly interpretable than the predefined variables in the pixel-share MNL.

For these reasons, the models developed in this thesis should be seen as exploratory tools for analysing stated pedestrian route preferences and visual route attractiveness. They should not be used as direct prediction tools for actual pedestrian flows or as stand-alone tools for making planning decisions. Future research should validate the findings using observed walking behaviour, improve the representation of full walking routes, examine specific visual street-level interventions, test whether the results are transferable to other contexts and population groups, and improve the explainability of image-based route choice models.

Overall, this thesis shows that pedestrian route choice is shaped by more than travel time alone. Visual street-level information can be incorporated into route choice models at different levels of complexity, and the importance of this information differs across walking purposes. The findings therefore support a broader view of pedestrian route choice, where walking routes are not only links between origins and destinations, but also street environments that pedestrians experience along the way.

Responsible Use of Generative AI

During the writing process of this thesis, generative artificial intelligence tools such as ChatGPT and Copilot were used as supportive tools for checking grammar, structure, and formulation. These tools were used to improve the clarity, flow, and readability of the text. They were also used to support the development of code and visual material where relevant. Sensitive or personally identifiable information was not shared with these tools. The research was conducted in accordance with the Human Research Ethics Committee procedures of TU Delft. The research design, methodological choices, data analysis, interpretation of results, and final conclusions remain the responsibility of the author. All AI-assisted output was critically reviewed, edited, and verified by the author before being included in the thesis. References, results, and interpretations were checked against the original data, literature, and model outputs.

Contents

Preface i

Summary ii

Table of Contents vii

List of Figures ix

List of Tables x

Abbreviations xi

1 Introduction 1
- 1.1 Pedestrian Route Choice Behaviour . . . 1
- 1.2 Pedestrian Modelling in Urban Mobility . . . 2
- 1.3 Computer Vision in Route Choice Modelling . . . 2
- 1.4 Research Gap . . . 3
- 1.5 Research Objective and Research Questions . . . 3
- 1.6 Thesis Outline . . . 3

2 Background and Related Work 4
- 2.1 Route Choice Modelling . . . 4
  - 2.1.1 Random Utility Theory and Discrete Choice Models . . . 4
  - 2.1.2 Challenges in Route Choice Modelling . . . 5
  - 2.1.3 Collecting data for Route Choice Modelling . . . 5
- 2.2 Pedestrian Behaviour and Route Choice . . . 6
  - 2.2.1 Behavioural Foundations of Pedestrian Route Choice . . . 6
  - 2.2.2 Factors Influencing Pedestrian Route Choice . . . 6
- 2.3 Computer Vision for Urban Analysis . . . 8
  - 2.3.1 Computer Vision Methods . . . 8
  - 2.3.2 Computer Vision in Discrete Choice Modelling . . . 9
- 2.4 Synthesis and Knowledge Gap . . . 10

3 Methodology 11
- 3.1 Research Design Overview . . . 11
- 3.2 Route Choice Attribute Selection . . . 13
  - 3.2.1 Trip Characteristics . . . 14
  - 3.2.2 Built- and Natural-Environment Factors . . . 14
  - 3.2.3 Socio-Demographic Factors . . . 14
  - 3.2.4 Contextual Factors . . . 15
- 3.3 Street-Level Image Data Collection . . . 15
  - 3.3.1 Pedestrian Network Extraction . . . 15
  - 3.3.2 Street-Level Image Sampling Strategy . . . 16
  - 3.3.3 Image Analysis and Filtering . . . 17
  - 3.3.4 Construction of Visual Categories . . . 18
- 3.4 Stated Preference Experiment . . . 20
  - 3.4.1 Stated Preference Approach . . . 20
  - 3.4.2 Survey Design . . . 21
  - 3.4.3 Choice Task Design . . . 21
  - 3.4.4 Survey Implementation . . . 24
- 3.5 Modelling Framework . . . 25
  - 3.5.1 Discrete Choice Modelling Framework . . . 25

3.5.2 Model 1: Baseline MNL 26
3.5.3 Model 2: Pixel-Share MNL 26
3.5.4 Model 3: Computer-Vision Enriched Discrete Choice Model 27
3.5.5 Model Estimation and Optimisation 28
3.6 Model Evaluation 29
3.6.1 Parameter Estimates and Relative Importance 30
3.6.2 Predictive Performance Metrics 30
3.6.3 Qualitative Model Validation 30
3.6.4 Spatial Application 31

4 Data Collection and Preparation 33
4.1 Survey Data Collection 33
4.2 Survey Data Cleaning 33
4.3 Image Data Preparation 34
4.4 Train-Test Split 34

5 Results 35
5.1 Descriptive Results 35
5.1.1 Socio-Demographics and Walking Habits 35
5.1.2 General Choice Task Patterns 36
5.1.3 Single-Image Perceptions 37
5.2 Model Estimation Results 38
5.2.1 Selected Model Settings 38
5.2.2 Parameter Estimates 39
5.2.3 Predictive Performance 41
5.3 Qualitative Model Validation 42
5.3.1 Comparison with Perception Scores 43
5.3.2 High- and Low-Utility Images 44
5.3.3 Same-Street Image Comparison 46
5.3.4 Image Augmentation Results 47
5.4 Spatial Application of the CV-DCM 48
5.4.1 Spatial Utility Distribution 48
5.4.2 Route Comparisons 50

6 Discussion 51
6.1 Interpretation of Main Findings 51
6.1.1 Prediction of Route Choice Preferences with Visual Information 51
6.1.2 Preference for Greener Walking Environments 52
6.1.3 Differences Between Trip Purposes 52
6.1.4 Differences Between Shortest Route and Preferred Routes 53
6.1.5 Model Interpretability 53
6.2 Strengths and Limitations 54
6.2.1 Comparison of Multiple Model Specifications 54
6.2.2 Trip-Purpose-Specific Modelling 54
6.2.3 Model Validation Beyond Model Fit 54
6.2.4 Sample Generalisability and Hypothetical Bias 55
6.2.5 Simplification of the Walking Environment 55
6.3 Research Implications 56
6.3.1 Pedestrian Route Choice Modelling 56
6.3.2 Route Attractiveness in Urban Planning and Policy 57
6.3.3 Artificial Intelligence in Behavioural Modelling and Policy Analysis 57
6.4 Recommendations and Future Research 58
6.4.1 Practical Recommendations 58
6.4.2 Future Research 59

7 Conclusion 61
7.1 Answer to the Research Questions 61
7.2 Main Conclusion 63

7.3 Contributions 63

A Factors Influencing Pedestrian Route Choice 68

B Street-Level Image Extraction and Processing 71
B.1 Step 1: Pedestrian Network Extraction 71
B.2 Step 2: Street-Level Image Sampling and Collection 72
B.3 Step 3: Image Analysis and Filtering 73
B.4 Step 4: Construction of Visual Categories 74

C Street-Level Images 75
C.1 Filtered Images 75
C.2 Visually Categorised Images 76

D Survey Pilot Design 79
D.1 Choice Task Mock Designs 79
D.2 Pilot Survey 80

E Final Survey Design 82
E.1 Survey Questions 82
E.2 Final Choice Task Efficient Design 86
E.2.1 Experimental Design 86
E.3 Pseudocode for Choice Task Design Construction 88

F Socio-Demographics 89

G Additional Results 91
G.1 Additional Choice Task Patterns 91
G.2 Alternative Pixel-Share Model Specifications 92
G.3 CV-DCM Hyperparameter Tuning Results 93
G.4 Same-Street Qualitative Validation Results 95
G.5 AI Image Augmentations 96
G.6 Spatial Application Route Comparisons 98

List of Figures

1 Overview of the three model specifications and the trade-off between interpretability and flexibility. . . . ii
2 Summary of the four trip-purpose contexts and the relative importance of travel time and visual information. . . . iii

2.1 Image feature extraction and classification, from van Cranenburgh and Garrido-Valenzuela (2025). . . . 9

3.1 Research design overview 12
3.2 Pedestrian networks for the three largest cities in the Netherlands 16
3.3 Spatial distribution of sampled Street View image locations in Amsterdam 17
3.4 Examples of accepted and rejected street-level images 18
3.5 Distributions of visual characteristics of filtered images 19
3.6 Example of a choice task in the context of walking home. 22
3.7 Cyclic assignment of respondents to design blocks 24
3.8 Rotation of choice task sets across trip-purpose contexts 24
3.9 Combined block assignment and context rotation across respondents 24
3.10 Model structure of the CV-DCM, as presented by van Cranenburgh and Garrido-Valenzuela (2025) 27

5.1 Main demographic characteristics of the respondent group. 35
5.2 General walking habits of the respondent group. 36
5.3 Chosen travel time and greenery level per trip purpose. 36
5.4 Spearman correlation matrix of perceived street characteristics. 37
5.5 Highest and lowest predicted image utilities by trip purpose. 44
5.6 Images with the largest rank differences across trip-purpose models. 45
5.7 Same-street comparison of predicted image utilities. 46
5.8 AI-edited image scenarios and changes in predicted CV-DCM image utility. 47
5.9 Distribution of AI-edited image utility changes per trip purpose for eight edited images. . . 47
5.10 Case study area in Amsterdam-Zuid used for the spatial application 48
5.11 Average street-segment image utility in Amsterdam-Zuid by trip-purpose model 49
5.12 Example street environments with lower and higher image utilities in the spatial application 49
5.13 Comparison between the shortest route and the highest-utility route from Vrije Universiteit Amsterdam to the Gijsbrecht van Aemstelpark by trip-purpose model. 50

C.1 Examples of images labelled using filtering thresholds 75
C.2 Examples of images classified by greenery presence levels 76
C.3 Examples of images classified by motorised vehicle presence 76
C.4 Examples of images classified by people presence 76
C.5 Combined classification of images by vehicle presence, greenery presence, and people presence (first 6 out of 12) 77
C.6 Combined classification of images by vehicle presence, greenery presence, and people presence (last 6 out of 12) 78

D.1 Overview of mock choice task designs used during the initial development phase 79
D.2 Pilot choice task design using crossing icons and trip-purpose context 80

E.1 Overview of choice task designs used in the final survey 83
E.2 Example perception question used in the final survey 84

F.1 Distribution of responses for the screening and demographic questions. 89
F.2 Distribution of responses for the background questions. 90

G.1 Distribution of chosen route attributes by trip purpose. 91
G.2 Same-street comparison of predicted image utilities for all trip purposes. 95
G.3 AI-edited image scenarios and changes in predicted CV-DCM image utility. 97
G.4 Comparison between the shortest-distance route and the highest-utility route for the selected origin-destination pairs in Amsterdam-Zuid. 98

List of Tables

3.1 Overview of selected attributes relevant for pedestrian route choice 13
3.2 Number of images per category after filtering 18
3.3 Counts for combined image characteristics 19
3.4 Attributes included in the final survey 21
3.5 Pilot estimates and adjusted priors for the final design 23

5.1 Selected hyperparameter configurations for the CV-DCM 39
5.2 Performance and parameter comparison across model types and trip purposes. 39
5.3 Contribution to the maximum possible utility difference in the Pixel-share MNL models. . 40
5.4 Contribution to the maximum possible utility difference in the CV-DCM models. 41
5.5 Average perception scores of the top and bottom 20 image utilities by model 43

B.1 City centre reference locations and coordinates used for each city 72
B.2 Rule-based image analysis and filtering thresholds 73
B.3 Binning methods used to construct visual categories 74

D.1 Image category definitions and prior utilities 81
D.2 Crossing types and prior utilities 81

E.1 Initial demographic questions 82
E.2 Final background and walking behaviour questions 85
E.3 Pilot estimates and adjusted priors for the final design 86
E.4 Image category definitions and priors for the image categories of the final survey 86
E.5 Final experimental design (Ngene output) 87

G.1 Performance and parameter comparison for the pixel-share MNL models. 92
G.2 Performance and parameter estimates for the pixel-share MNL models with an alternative-specific constant. 93
G.3 Overview of the CV-DCM hyperparameter search procedure. 93
G.4 Selected CV-DCM hyperparameter configurations based on validation cross-entropy. . . 93
G.5 Performance and hyperparameter configuration for the trip-purpose-specific CV-DCM models. 94

Abbreviations and Acronyms

AI	Artificial Intelligence
API	Application Programming Interface
ASC	Alternative-Specific Constant
BIC	Bayesian Information Criterion
CE	Cross-Entropy
CV	Computer Vision
CV-DCM	Computer Vision-Enriched Discrete Choice Model
DCM	Discrete Choice Model
EPA	Engineering and Policy Analysis
GIS	Geographic Information Systems
GPS	Global Positioning System
HR	Hit-Rate
LL	Log-Likelihood
MNL	Multinomial Logit
RP	Revealed Preference
RUM	Random Utility Maximisation
RUT	Random Utility Theory
SGD	Stochastic Gradient Descent
SP	Stated Preference
SUMS	Sustainable Urban Mobility and Space
TU Delft	Delft University of Technology
XAI	Explainable Artificial Intelligence

1

Introduction

Walking plays an important role in everyday mobility in the Netherlands. In 2023, walking accounted for about a quarter of all trips in the Netherlands, which is close to the share of cycling (Kennisinstituut voor Mobiliteitsbeleid, 2024; Centraal Bureau voor de Statistiek, 2024). Many transport models therefore incorporate pedestrian behaviour. However, this behaviour is often simplified, for example by assuming that pedestrians always choose the shortest or fastest route (Prato, 2009; Zafri and Sevtsuk, 2026). This assumption misses important parts of pedestrian route choice, since pedestrians may also consider environmental aspects when choosing their routes (Basu et al., 2022; Tong and Bode, 2022; López-Lambas et al., 2021; Sevtsuk et al., 2021).

For urban policy, this simplification matters. Walking is increasingly linked to sustainable mobility, public health, and more efficient use of public space (European Commission, 2021; Fernández Aguilar et al., 2023; Ministry of Infrastructure and Water Management, 2022; Kennisinstituut voor Mobiliteitsbeleid, 2024). This is particularly relevant in Dutch cities, where population growth, combined with limited space for new infrastructure, puts pressure on the mobility system and this pressure is expected to further increase in the future (Ministry of Infrastructure and Water Management, 2022; Kennisinstituut voor Mobiliteitsbeleid, 2024). Transport models are widely used to support mobility planning and policy evaluation (Batty, 2009; Borchers et al., 2024). However, when these models represent pedestrian route choice in a simplified way, policy evaluations may provide an incomplete picture of how pedestrians respond to changes in the walking environment (Zafri and Sevtsuk, 2026). This is especially relevant for policies that aim to increase walkability or route attractiveness, because these aspects are not fully captured by distance or travel time alone (Basu et al., 2022; Sevtsuk et al., 2021).

1.1. Pedestrian Route Choice Behaviour

When pedestrians walk through a city, they usually have multiple possible paths between an origin and a destination. The choice between these paths is not only influenced by distance or travel time, but also by how streets are experienced and perceived (Basu et al., 2022; Tong and Bode, 2022). Visual street-level characteristics, such as traffic, sidewalk quality, greenery, and the presence of other people, may influence which route pedestrians prefer (Basu et al., 2022; Choi and Kang, 2025; Sevtsuk et al., 2021). Pedestrians may prefer a street that feels safer, looks greener, or is less busy, even when it is not part of the shortest route (López-Lambas et al., 2021). These preferences may also differ between walking contexts, for example when walking to public transport, to work or school, home, or in free time (Agrawal et al., 2008; Cerin et al., 2007; Yang and Diez Roux, 2012).

In research on pedestrian route choice, two perspectives are particularly relevant. Walkability and perception research focuses on how pedestrians experience and evaluate their environment, often using visual and environmental characteristics to describe how attractive, safe, or comfortable a street feels (Choi and Kang, 2025; Rossetti et al., 2019). Route choice modelling focuses on how people choose between alternative routes and how route characteristics influence these choices (Prato, 2009; Basu et al., 2022). In route choice models, routes are usually represented using structured variables such as distance, travel time, network characteristics, or GIS-based built-environment data. These variables are measurable and interpretable, but may not fully reflect how pedestrians perceive the walking environment at street level (Hoogendoorn and Bovy, 2004). At the same time, walkability and perception studies give insight into the visual and behavioural experience of walking, but do not always show how these experiences influence choices between route alternatives. This suggests a need to connect structured route choice modelling with more visual and behaviour-oriented approaches from walkability research. This connection is especially relevant for pedestrian models used in urban planning, because route attractiveness depends on more than distance or travel time alone.

1.2. Pedestrian Modelling in Urban Mobility

Transport models are widely used to support urban mobility planning and policy evaluation. For example, they can be used to explore how infrastructure changes or policy interventions may affect mobility systems before they are implemented (Batty, 2009; Borchers et al., 2024). Although pedestrians are an important part of urban mobility, transport modelling has historically focused more on motorised traffic than on walking (Zafri and Sevtsuk, 2026). As a result, pedestrian modelling remains less developed than modelling for motorised transport modes. This creates opportunities to further develop pedestrian models for estimating walking flows, identifying important walking routes, and assessing how changes in the built environment may influence pedestrian activity (Zafri and Sevtsuk, 2026).

Different pedestrian modelling approaches exist, ranging from shortest-path assignment to network-based flow models, agent-based simulations, and Discrete Choice Models (DCMs) (Prato, 2009; Zafri and Sevtsuk, 2026). However, pedestrian behaviour is often simplified to keep models practical and computationally manageable. In route choice modelling, this often means relying on shortest-path assumptions (Prato, 2009). These assumptions are useful for implementation, but limited when models are used to evaluate changes in the quality of the walking environment. This shows the relevance of modelling approaches that remain practical for urban mobility applications, while better representing pedestrian route choice preferences. One way to do this is to incorporate visual information in pedestrian route choice models.

1.3. Computer Vision in Route Choice Modelling

Recent developments in Computer Vision (CV) make it possible to use street-level images to describe the visual environment from a pedestrian perspective. By analysing street-level images, visual characteristics of streets can be represented numerically and linked to how people experience urban environments (Dubey et al., 2016). Several studies have used CV methods to extract visual features such as greenery or crowdedness from street-level images and to analyse how these features relate to perceived safety, comfort, or attractiveness (Choi and Kang, 2025; Rossetti et al., 2019).

The relevance of this type of visual information for pedestrian route choice has been shown by Sevtsuk et al. (2021). Their study shows that street-level characteristics derived from images can help explain observed walking route choices within a DCM framework. However, this approach relies on a pre-defined set of visual characteristics, which are first measured from the images and then included as variables in the route choice model.

More recent work combines CV more directly with choice modelling through the Computer Vision-Enriched Discrete Choice Model (CV-DCM), introduced by van Cranenburgh and Garrido-Valenzuela (2025). The CV-DCM builds on the discrete choice modelling framework, but instead of selecting specific visual characteristics in advance, it uses a deep neural network to create a numerical representation of the full street-level image. This representation can be combined directly with non-visual attributes to explain choices between alternatives. The CV-DCM has been applied to cycling route choice in Rotterdam by Terra et al. (2025), showing that this approach can be used in a route choice setting where images represent the visual environment of route alternatives. Since pedestrian route choice also involves routes that differ in their visual street-level characteristics, the CV-DCM provides a suitable framework for studying pedestrian route choice preferences.

1.4. Research Gap

Pedestrian route choice is shaped by non-visual attributes, such as distance and travel time, but also by visual street-level characteristics and contextual factors, including greenery, traffic, sidewalk quality, the presence of other people, and trip purpose (Basu et al., 2022; Tong and Bode, 2022; Hoogendoorn and Bovy, 2004; Sevtsuk et al., 2021). However, pedestrian route choice models often still rely on shortest-path assumptions or predefined route attributes, which may not fully reflect how pedestrians perceive and evaluate the walking environment (Prato, 2009; Hoogendoorn and Bovy, 2004; Zafri and Sevtsuk, 2026). Although walkability, perception, and computer vision studies show that street-level visual characteristics are related to how streets are experienced, it remains unclear how and to what extent visual information should be incorporated into pedestrian route choice models and how its role differs across walking contexts (Choi and Kang, 2025; Rossetti et al., 2019; Dubey et al., 2016; Basu et al., 2022).

1.5. Research Objective and Research Questions

The objective of this thesis is to address the research gap by examining how visual and non-visual route attributes influence pedestrian route choice preferences in the Netherlands, using discrete choice models that compare different levels of visual information across trip-purpose contexts.

The main research question therefore is:

To what extent do visual and non-visual route attributes influence pedestrian route choice preferences in the Netherlands?

To answer the main research question, the following sub-questions are defined:

1. Which visual and non-visual route attributes are relevant to pedestrians when choosing a route?
2. How can the trade-offs pedestrians make between visual and non-visual route attributes be measured?
3. How can different levels of visual information be incorporated into discrete choice models for pedestrian route choice preferences?
4. How does the inclusion of visual information at different levels affect model performance and the interpretation of pedestrian route choice preferences across trip-purpose contexts?
5. How can a visually enriched route choice model be applied to analyse trip-purpose-specific route preferences across a pedestrian network?

Together, these sub-questions structure the analysis needed to answer the main research question. The first sub-question identifies the relevant visual and non-visual route attributes and defines the scope of the analysis. The second explains how the trade-off between these attributes can be measured and supports the creation of the dataset. The third shows how different levels of visual information can be incorporated into discrete choice model formulations. The fourth evaluates how these model specifications affect performance and interpretation across trip-purpose contexts. Finally, the fifth question moves from model estimation to spatial application by showing how an image-based route choice model can be used to examine differences in route preferences across a pedestrian network.

This thesis is conducted in collaboration with the SUMS department at TNO. The research connects pedestrian route choice modelling with urban mobility analysis and policy evaluation. This makes the topic relevant for the Engineering and Policy Analysis (EPA) programme, because it combines behavioural modelling, spatial analysis, and machine learning to study a socio-technical mobility problem.

1.6. Thesis Outline

This thesis is structured as follows. Chapter 2 presents the literature review and discusses pedestrian route choice behaviour, route choice modelling, and computer vision for visual street-level representation. Chapter 3 describes the methodology, including the Stated Preference (SP) experiment, model specifications, and model evaluation. Chapter 4 describes the data collection, cleaning, and preparation process. Chapter 5 presents the model results, qualitative validation, and spatial application. Chapter 6 discusses the findings, limitations, implications, and recommendations for future research. Finally, Chapter 7 concludes the thesis and summarises its main contributions.

2

Background and Related Work

This chapter reviews the literature relevant to modelling pedestrian route choice behaviour and the role of visual street-level characteristics. The aim is to provide the theoretical foundation for understanding how pedestrians choose routes, how these choices can be represented in a discrete choice modelling framework, and how visual information from street-level images can be included in this process.

The chapter is structured around three main components. First, Section 2.1 introduces route choice modelling, with a focus on discrete choice models and the methodological challenges related to them. Second, Section 2.2 discusses pedestrian route choice behaviour and explains why walking routes are influenced by both objective route attributes and subjective perceptions of the walking environment. It also discusses which attributes are most often considered in the literature. Third, Section 2.3 examines how computer vision can be used to represent visual street-level information and how this information can be incorporated into route choice models. Finally, Section 2.4 summarises the main takeaways from the literature and identifies the main gaps addressed in this thesis.

2.1. Route Choice Modelling

Route choice modelling aims to represent how people select paths between origins and destinations. It focuses on how individuals choose a route based on the available alternatives, their preferences, and their perceptions (Bovy and Stern, 1990). Traditional route choice models often relied on deterministic shortest-path or minimum travel time approaches to approximate this behaviour (Prashker and Bekhor, 2004; Prato, 2009). These approaches assume that travellers have full knowledge of the network and always choose the lowest-cost alternative. While this is computationally efficient, it may not fully reflect real-world behaviour. As a result, route choice research has moved towards probabilistic models that can account for perceived costs, uncertainty, and variation in individual decision-making (Prashker and Bekhor, 2004; Prato, 2009).

2.1.1. Random Utility Theory and Discrete Choice Models

To model the complexity of route choice, Random Utility Theory (RUT) has become a widely used theoretical framework. In this theory, utility is used as a numerical value that represents the relative preference of an individual for different alternatives. This makes it possible to rank the available routes (McFadden, 1974; Ben-Akiva and Lerman, 1985). RUT assumes that utility consists of an observable systematic part and a random unobservable part. This reflects that not all factors influencing a choice can be explicitly included in the model. When individuals are assumed to choose the alternative that gives them the highest utility, the model follows the logic of a Random Utility Maximisation (RUM) (McFadden, 1974). DCMs use this utility-based framework to calculate the probability of choosing between a finite set of alternatives. They represent route choice as a set of trade-offs between different route attributes through an explicit utility function (Ben-Akiva and Lerman, 1985). This makes it possible to show how people balance travel time with other factors, such as comfort or safety, while keeping the model interpretable.

In addition to DCMs, several other approaches are used to model pedestrian route choice. Network-based models are used to estimate pedestrian flows and identify important walking routes within a network (Zafri and Sevtsuk, 2026). Rule-based models represent route choice using simplified decision rules, such as minimising distance or avoiding specific road types (Prashker and Bekhor, 2004). More recently, data-driven methods, such as machine learning, have been applied to route choice modelling. These methods can capture complex patterns in data and may perform well in predicting route choices, but they often provide limited insight into the underlying decision-making process. They also make it more difficult to interpret the influence of individual route attributes (Prato, 2009). Since this thesis aims to examine how visual and non-visual route attributes influence pedestrian route choice preferences, rather than only predict which route is chosen, DCMs provide a suitable and theoretically grounded framework.

2.1.2. Challenges in Route Choice Modelling

There are several challenges when applying DCMs to route choice. One issue is route overlap, which is relevant when many route alternatives share parts of the same roads. The Multinomial Logit (MNL) model, which is often used as a baseline model, assumes that the unobserved utility components are independently and identically distributed according to an extreme value distribution (McFadden, 1974; Ben-Akiva and Lerman, 1985). This means that routes are treated as fully independent from each other. In the case of overlapping routes, this assumption may become unrealistic, because routes that share many of the same streets are likely to be perceived as similar. This can lead to an overestimation of the combined choice probability of similar overlapping routes, because the MNL model does not account for the correlation between alternatives that share parts of the same path (Prato, 2009). Several models that extend the MNL can be used to address this, such as path-size or C-logit models (Prashker and Bekhor, 2004). These models account for the degree of overlap within their utility functions.

Another issue in route choice modelling is generating realistic choice sets. In reality, the number of possible routes between an origin and a destination can become very large, especially in a dense network. However, people rarely consider all possible routes between an origin and a destination (Prato, 2009). This means that generating choice sets is about approximating the set of routes that people are likely to consider. Including unrealistic routes may bias parameter estimates, while excluding relevant routes may lead to incorrect predictions. Methods such as k-shortest path or link-cost simulation are often used to generate route alternatives (Bekhor et al., 2006; Prato, 2009). An alternative approach is the recursive logit model introduced by Fosgerau et al. (2013), which models route choice as a series of link-based decisions instead of requiring a predefined set of route alternatives.

2.1.3. Collecting data for Route Choice Modelling

To estimate DCMs, data on route choices is needed. Because underlying preferences cannot be observed directly, they must be inferred from the choices people make. In transport modelling, these preferences are usually measured using Revealed Preference (RP) or Stated Preference (SP) approaches (Ben-Akiva and Lerman, 1985; Train, 2009). RP data is based on observed real-world behaviour, for example by analysing GPS traces, walk-alongs, or video-recorded route choices (Duives, 2025; Sevtsuk et al., 2021). This makes it possible to study actual decisions in real-world conditions. At the same time, it is difficult to vary specific route attributes in a controlled way, since the observed choices depend on real-world situations.

SP data is collected through hypothetical choice experiments in which respondents choose between predefined alternatives. In these experiments, route attributes can be controlled and varied systematically (Ben-Akiva and Lerman, 1985). This allows researchers to isolate the effect of specific factors. Because SP responses are based on stated intentions rather than actual behaviour, the results may be affected by hypothetical bias or strategic answering (Ben-Akiva and Lerman, 1985; Brownstone et al., 2000).

2.2. Pedestrian Behaviour and Route Choice

Pedestrians differ from other travellers because they move at a slower pace and are directly exposed to their surroundings. As a result, walking behaviour is flexible and can be influenced by detailed street-level characteristics (Tong and Bode, 2022; Duives, 2025). Greenery, traffic, sidewalk quality, and lighting can affect how attractive, safe, or comfortable a street is perceived to be. Perception and walkability studies provide evidence that such visual and environmental characteristics are related to perceived street quality (Choi and Kang, 2025). However, perceived attractiveness does not automatically imply that pedestrians will deviate from the shortest path in actual route choice. Route choice studies show that pedestrians may accept longer routes when alternatives offer a more attractive or comfortable walking environment (López-Lambas et al., 2021; Sevtsuk et al., 2021). A systematic review by Basu et al. (2022) also shows that built-environment factors, trip characteristics, and socio-demographic characteristics are associated with pedestrian route choice, while route familiarity and trip purpose shape how these attributes are perceived and evaluated.

2.2.1. Behavioural Foundations of Pedestrian Route Choice

The theoretical foundation of pedestrian route choice can be found in wayfinding and spatial cognition research. These fields assume that pedestrians do not have complete knowledge of the transport network. Instead, they form mental representations based on what they see, know, and experience (Golledge, 1992; Bovy and Stern, 1990). Therefore, route choice is not only based on objective distance or travel time, but also on how routes are perceived in terms of comfort, safety, and attractiveness.

Behavioural frameworks describe pedestrian route choice as part of a broader decision process. Hoogendoorn and Bovy (2004) distinguish between strategic, tactical, and operational levels of pedestrian behaviour. Route choice mainly takes place at the tactical level. At this level, pedestrians evaluate route alternatives based on perceived attributes and choose the alternative with the lowest subjective disutility. This behavioural interpretation fits utility-based route choice models, in which route preferences are represented as trade-offs between multiple route attributes. However, route choice takes place within a wider walking context. Broader decisions, such as the purpose of the trip, are already part of the situation in which the route is evaluated. This is also highlighted by Duives (2025), who places route choice within a broader set of travel decisions, including activity choice, departure time choice, and mode choice.

Pedestrian route choice can therefore be understood as a process in which route alternatives are evaluated within a specific walking context. Tong and Bode (2022) describe this as a sequential process in which pedestrians perceive their surroundings, interpret this information, evaluate route attributes, and then make a choice. In this view, utility is determined not only by objective route characteristics, but also by how these characteristics are perceived by the pedestrian. This helps explain why pedestrians may not evaluate attributes such as travel time, traffic, greenery, and sidewalk quality separately, but rather as part of the overall perception of a route.

2.2.2. Factors Influencing Pedestrian Route Choice

The described behavioural frameworks described show that pedestrian route choice is influenced by a combination of objective and subjective factors. In their systematic review, Basu et al. (2022) identify 105 factors associated with pedestrian route choice and group them into three broad categories: trip characteristics, built- and natural-environment factors, and socio-demographic factors. This section uses these categories to structure the discussion of factors influencing pedestrian route choice. Contextual factors are added as a separate group, because they can influence how route attributes are perceived and evaluated in different walking situations. A complete overview of the factors identified by Basu et al. (2022) and their reported directions of influence is provided in Appendix A.

Trip Characteristics

Trip characteristics describe the practical conditions of the route and the trip. Distance and travel time are among the most commonly studied factors in pedestrian route choice research. Pedestrians generally prefer routes that are shorter and faster (Basu et al., 2022; Prato, 2009). However, pedestrians do not always choose the shortest route. Several studies show that pedestrians may accept a longer route when the alternative offers a more attractive or comfortable walking environment (López-Lambas et al., 2021; Sevtsuk et al., 2021).

In addition to distance and travel time, factors such as traffic and waiting times at crossings also influence route attractiveness (Basu et al., 2022; Tong and Bode, 2022). From a behavioural perspective, these attributes affect the perceived effort of taking a certain route (Duives, 2025).

Built- and Natural-Environment Factors

The built- and natural-environment includes the physical characteristics of the street and its surroundings. These characteristics influence walking behaviour through their effects on comfort, safety, and accessibility. Together, these aspects shape the walkability of a street (Duives, 2025). Research shows that pedestrian-specific infrastructure, such as sidewalk width and maintenance, can influence route choice (Basu et al., 2022; López-Lambas et al., 2021). In addition, features such as greenery, open spaces, and active shop windows are associated with higher route attractiveness, while higher traffic volumes are generally linked to lower perceived safety and lower route attractiveness (Choi and Kang, 2025; Basu et al., 2022).

Many of these environmental characteristics are experienced directly at street level, because pedestrians perceive and interpret their surroundings throughout the walk (Tong and Bode, 2022). At the same time, not all relevant environmental factors are visual. Other factors, such as weather, noise, and smell also influence how streets are perceived and can affect walking comfort (Basu et al., 2022; Tong and Bode, 2022). This shows that visual street-level characteristics are important for pedestrian route choice, but do not capture the full walking experience.

Socio-Demographic Factors

Pedestrian route choice varies across individuals and contexts. At a larger scale, cultural and spatial factors influence how pedestrians evaluate their environment. This means that walking behaviour can differ across countries and urban settings. Levine and Norenzayan (1999) show that average walking speeds differ between countries and can be related to climate, economic conditions, and cultural values. In addition, the effect of built-environment characteristics on walking behaviour can differ across contexts and countries (Lin et al., 2015). This suggests that the environments people are familiar with can influence how they perceive other routes and street environments.

At a more individual level, personal characteristics such as age, gender, and income are often included in route choice studies. Basu et al. (2022) identify age and gender as the most frequently studied demographic factors. For example, a study in the Netherlands shows that older pedestrians prefer routes with better environmental quality, lighting, and pavement conditions (Borst et al., 2008). Differences related to gender are often discussed in relation to safety and comfort, where women may evaluate routes differently than men (Mazzulla et al., 2024). However, the review by Basu et al. (2022) shows that these effects are not always consistent across the literature.

Contextual Factors

In addition to route and personal characteristics, contextual factors can influence how pedestrians evaluate route alternatives. One important contextual factor is route familiarity. Pedestrians often prefer routes they already know, which means that route choice can be influenced by habits and previous experiences (Duives, 2025). Familiarity can also influence how the walking environment is perceived, because a pedestrian who knows an area may already know which streets, crossings, or shortcuts they prefer.

Trip purpose is also discussed in the literature as an important context for walking behaviour. Walking behaviour has been shown to differ by trip purpose, including differences in walking distance, walking duration, and the importance of the built environment (Larranaga and Cybis, 2014; Hatamzadeh et al., 2014; Yang and Diez Roux, 2012; Watson et al., 2021; Mondschein, 2018). For example, Larranaga and Cybis (2014) find that the built environment plays a stronger role for walking trips related to shopping, recreation, health, and personal business than for more goal-oriented trips such as work or study. For walking trips to public transport, Agrawal et al. (2008) find that pedestrians often prioritise factors such as travel time, distance, safety, and waiting time at crossings. These findings show that trip purpose does not describe the route itself, but the context in which route alternatives are evaluated.

2.3. Computer Vision for Urban Analysis

Computer Vision (CV) is increasingly used in urban research to analyse street-level images and to describe the visual environment. Using street-level images, studies have extracted visual characteristics such as greenery, enclosure, visible sky, buildings, road space, and traffic presence (Dubey et al., 2016; Ma et al., 2021; Rossetti et al., 2019). These studies show that CV can make visual qualities of urban environments measurable. However, their focus is mainly on the perceptions of the environment, not how these perceptions affect route choice.

Route choice studies have traditionally relied on GIS-based variables, such as land use, street connectivity, and neighbourhood density (Prato, 2009; Basu et al., 2022). These variables describe the general structure of the built environment, but they do not directly represent the street-level visual experience of pedestrians moving through the city. Studies that combine street-level image information with pedestrian route choice are still limited. However, they suggest that visual characteristics can add information that is not captured by traditional route attributes (Sevtsuk et al., 2021). This suggests that CV can help connect the visual qualities studied in urban perception research with the structured models of pedestrian route choice.

2.3.1. Computer Vision Methods

In urban research, several types of CV methods are used to extract information from street-level images. In general, CV focuses on extracting meaningful information from images and other forms of visual data. Images consist of pixels and each pixel has a location and contains colour information, usually represented by the amount of red, green, and blue. This means that images can be represented as numerical data. However, using all individual pixels directly is often inefficient, because an image contains a large amount of data and single pixels have limited meaning on their own. Therefore, CV methods are used to extract more useful information from images.

One commonly used approach is semantic segmentation. This method classifies each pixel into predefined visual categories, such as vegetation, buildings, sky, sidewalks, roads, or vehicles. This allows the image to be represented by interpretable variables. For example, the percentage of pixels representing trees, grass or visible sky can be calculated and used to describe environmental characteristics such as greenness or openness. These characteristics can then be linked to urban quality and walkability (Rossetti et al., 2019; Ma et al., 2021; Sevtsuk et al., 2021; Choi and Kang, 2025).

A related approach is object detection, which identifies separate objects in an image, such as pedestrians, bicycles, or vehicles. These objects can be counted and used as proxies for activity levels or traffic conditions (Choi and Kang, 2025). Object detection focuses on individual elements in the street environment rather than on the overall visual composition of the image.

A third approach is deep feature extraction. Instead of relying on predefined categories, this method uses neural networks to learn visual patterns directly from images. A feature extractor transforms an image into a set of numerical features that describe the overall visual structure. These features are often referred to as embeddings or feature maps. They can be used to make predictions about the image or to use the image as input for another task. This process is shown in Figure 2.1. Feature maps can capture broader visual characteristics because they are based on the full image (Dubey et al., 2016; Choi and Kang, 2025). However, they are less interpretable, because they are numerical descriptions of many connected elements of the image rather than clearly defined visual attributes.

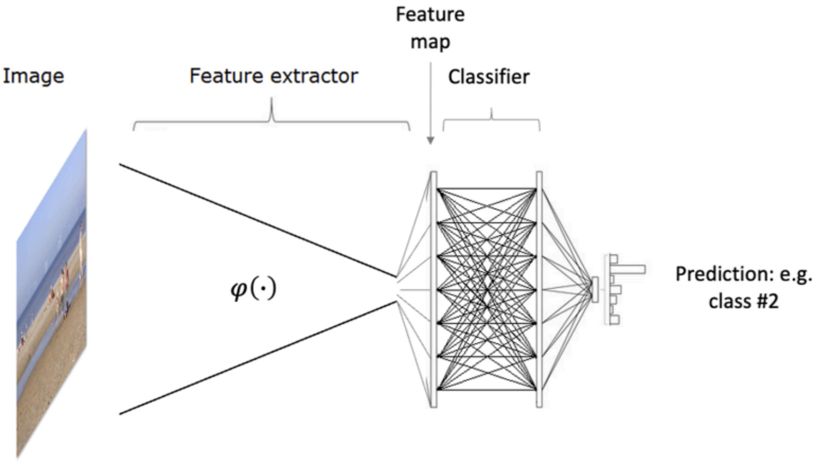


Figure 2.1: Image feature extraction and classification, from van Cranenburgh and Garrido-Valenzuela (2025).

2.3.2. Computer Vision in Discrete Choice Modelling

The use of visual information in discrete choice modelling is still relatively new, but recent studies show how street-level images can be connected to choice models. Rossetti et al. (2019) use semantic segmentation outputs as explanatory variables in discrete choice models of street perception. In their study, respondents choose between images based on qualities such as beauty, safety, or liveliness. Although this does not involve route choice, it shows that image-derived visual information can be included in discrete choice frameworks in a meaningful way.

In pedestrian route choice research, Sevtsuk et al. (2021) show that predefined visual characteristics extracted from street-level images can help explain observed walking choices. They used RP data in the form of anonymised GPS trajectories and looked at possible alternative routes to estimate a route choice model that includes both traditional route attributes and selected visual characteristics derived from Google Street View images. This study connects street-level visual information to actual walking behaviour. At the same time, the analysis is limited to the visual characteristics that were defined in advance. The visual information was included through variables that were selected before the model was estimated. This means that the model can only test the role of these predefined visual characteristics, while other visual information in the images is not included.

The Computer Vision-Enriched Discrete Choice Model (CV-DCM), introduced by van Cranenburgh and Garrido-Valenzuela (2025), provides a more direct way to include visual information in discrete choice models. Instead of first reducing images to predefined visual characteristics, the CV-DCM includes street-level images directly as part of the utility function. In this framework, a deep image encoder transforms images into numerical representations that are estimated together with the preference parameters. This allows the street environment to be represented using full images alongside non-visual attributes, such as distance or travel time, while remaining consistent with RUT. By including the full image, the CV-DCM may capture broader visual qualities, such as the overall appearance of a street.

The CV-DCM has been applied in residential location choice and cycling route choice (van Cranenburgh and Garrido-Valenzuela, 2025; Terra et al., 2025). In the cycling application, street-level images are used in a stated choice experiment to represent the cycling environment and to elicit route preferences. These images are then used within the CV-DCM to estimate utilities for different environments, together with other non-visual route attributes. This shows that the CV-DCM can be used in a route choice context.

2.4. Synthesis and Knowledge Gap

The reviewed literature shows that pedestrian route choice is influenced by a combination of objective route attributes, perceptions of the walking environment, and contextual factors. At the same time, current modelling approaches only partly reflect this complexity. Based on the literature discussed in this chapter, three main gaps can be identified.

The first gap is that many pedestrian route choice models still rely on shortest-path assumptions, rather than representing the broader set of factors identified in behavioural research. Pedestrian behaviour research shows that pedestrians evaluate routes based on comfort, safety, attractiveness, and the walking environment (Tong and Bode, 2022; Basu et al., 2022; Prato, 2009). This simplification limits how well models can represent the trade-offs pedestrians may make between distance, travel time, and the quality of the walking environment. As walking becomes more important in dense urban areas, this gap becomes increasingly relevant. Models that mainly focus on distance and travel time may be less suitable for evaluating pedestrian-oriented policies and street design.

The second gap is that the role of trip purpose is often not represented in pedestrian route choice models. Behavioural frameworks describe trip purpose as a strategic-level factor that shapes how route alternatives are evaluated (Hoogendoorn and Bovy, 2004). Studies show that walking behaviour differs by trip purpose, including differences in walking distance, walking duration, and sensitivity to environmental characteristics (Larranaga and Cybis, 2014; Yang and Diez Roux, 2012; Watson et al., 2021). However, many pedestrian route choice models either ignore trip purpose or assume that route choice preferences are the same across all walking trips. This means that there is limited insight into how route choice preferences may differ between more goal-oriented trips and less time-constrained trips.

The third gap is that it is still unclear what the added value is of using different levels of visual information in pedestrian route choice modelling. Walkability and urban perception research shows that visual street-level characteristics influence how pedestrians experience streets, while route choice models usually rely on structured or GIS-based route attributes (Choi and Kang, 2025; Rossetti et al., 2019). Recent work shows that visual information from street-level images can help explain observed pedestrian route choice, but this work relies on a predefined set of visual characteristics (Sevtsuk et al., 2021). The CV-DCM provides a way to include full street-level images directly in a discrete choice model and has been applied in residential location choice and cycling route choice (van Cranenburgh and Garrido-Valenzuela, 2025; Terra et al., 2025). However, it remains unclear whether more complex full-image representations provide additional explanatory value for pedestrian route choice preferences beyond predefined visual characteristics or non-visual route attributes.

Together, these gaps show that pedestrian route choice models do not yet fully account for the visual quality of routes or the purpose of the walking trip when representing route choice preferences. This thesis addresses these gaps by applying a computer vision-enriched discrete choice modelling approach to pedestrian route choice preferences in the Netherlands. It examines how visual and non-visual route attributes are reflected in route utility, and whether these preferences differ across trip-purpose contexts.

3

Methodology

This chapter describes the methodological approach used to answer the research questions of this thesis. It addresses the first three sub-questions: *Which visual and non-visual route attributes are relevant to pedestrians when choosing a route? How can the trade-offs pedestrians make between visual and non-visual route attributes be measured?* and *How can these attributes be incorporated into a discrete choice model for pedestrian route choice preferences?* The chapter also describes how the models are evaluated and how the image-based model is applied in an exploratory spatial analysis.

Section 3.1 presents a high-level overview of the research design. Section 3.2 describes the identification and selection of relevant route attributes. Section 3.3 explains the collection and processing of street-level images. Section 3.4 describes how the experiment is implemented. Section 3.5 introduces the modelling framework and the different models used to estimate pedestrian route choice preferences. Finally, Section 3.6 describes the methods used to evaluate model performance and behavioural consistency.

3.1. Research Design Overview

The aim of the research design is to examine how visual and non-visual route attributes influence pedestrian route choice preferences, and how these preferences differ across trip-purpose contexts. To do this, the research combines a Stated Preference (SP) experiment with street-level images and discrete choice modelling. This design is chosen because pedestrian route choice preferences are not only influenced by travel time, but also by how the walking environment is perceived. Since visual street-level characteristics are difficult to represent using only textual or numerical variables, street-level images are used to show the walking environment from a pedestrian perspective.

The research design uses three model types to compare different ways of incorporating route attributes into a Discrete Choice Model (DCM). A basic Multinomial Logit (MNL) model is used as a baseline model with only travel time. A pixel-share MNL model is used to include predefined visual attributes extracted from the street-level images. Finally, the Computer Vision-Enriched Discrete Choice Model (CV-DCM), introduced by van Cranenburgh and Garrido-Valenzuela (2025), is used to include full street-level images directly in the utility function. The CV-DCM is used because it can learn image representations from the complete image, instead of relying only on visual attributes that are defined beforehand.

The study follows a structured sequential workflow in which the different steps build on each other. An overview of the research design is shown in Figure 3.1.

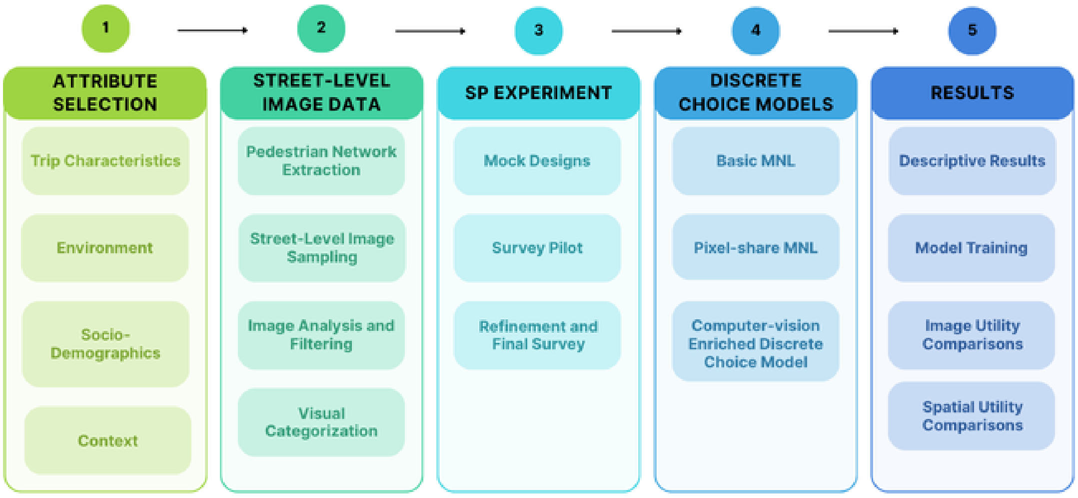


Figure 3.1: Research design overview

First, the relevant pedestrian route choice attributes are identified in Section 3.2. These attributes are based on the literature presented in Chapter 2 and on the list of attributes identified by Basu et al. (2022), as shown in Appendix A. Because pedestrian behaviour differs from other transport modes, the set of attributes is defined specifically for pedestrians. The selection includes trip characteristics, built- and natural-environment characteristics, socio-demographic factors, and contextual factors. From this broader set, a smaller set of attributes is selected for the experiment. A distinction is made between attributes included directly in the choice tasks, visual attributes represented through images, fixed contextual factors, and attributes collected separately through general survey questions.

Second, a dataset of street-level images is constructed in Section 3.3. The ten largest cities in the Netherlands are used to capture variation in urban environments and street characteristics. For each city, the pedestrian network is extracted and used to sample street locations. At these locations, street-level images are collected and filtered based on relevance and quality. The images are then analysed using semantic segmentation and object detection to extract visual features. These features are used to describe the selected visual attributes and to construct the image categories used in the experiment. A pseudocode description of this process is provided in Appendix B.

Third, an SP experiment is designed to collect data on pedestrian route choice preferences, as described in Section 3.4. The experiment is implemented as an online survey using PixelSurvey (Garrido-Valenzuela et al., 2025). In the choice tasks, respondents choose between two route alternatives that differ in travel time and visual street-level characteristics. These choices are made across different trip-purpose contexts, so that differences between walking contexts can be analysed. In addition to the choice tasks, the survey includes perception questions based on single images and general questions about walking habits and socio-demographic characteristics. The survey design is developed through several mock designs, a pilot survey, and a refined final design. This iterative process is used to improve the clarity of the choice tasks and reduce potential bias in the responses (Bliemer and Rose, 2024).

Fourth, the collected survey data is used to estimate multiple DCMs, which are formally described in Section 3.5. The basic MNL model is estimated first and serves as a baseline. The pixel-share MNL model then includes predefined visual variables derived from the street-level images. The CV-DCM combines travel time with full street-level images in the utility specification. Comparing these models makes it possible to evaluate whether visual information improves the explanation of stated pedestrian route choice preferences, and whether full images provide additional explanatory value compared with predefined visual variables.

Finally, the estimated models are evaluated and applied, as described in Section 3.6. Model performance is assessed using predictive performance metrics. The model outcomes are then further interpreted through parameter estimates, relative importance measures, and qualitative validation of image utilities. The trip-purpose-specific CV-DCM models are then applied to a pedestrian network in Amsterdam-Zuid to explore how image utilities can be represented spatially and how preferred routes differ from the shortest route. This final step does not aim to predict actual pedestrian flows. Instead, it demonstrates how the estimated model can be used to analyse spatial differences in route attractiveness across a pedestrian network.

3.2. Route Choice Attribute Selection

This part of the methodology addresses the first sub-question: *Which visual and non-visual route attributes are relevant to pedestrians when choosing a route?* It identifies the broader set of relevant factors from the literature and explains which attributes are selected for the stated preference experiment.

As described in Chapter 2, pedestrian route choice is influenced by many factors related to trip characteristics, the built and natural environment, socio-demographics, and the wider walking context (Basu et al., 2022; Tong and Bode, 2022; Duives, 2025). However, not all of these factors can be directly included in the experimental design. Including all 105 factors identified by Basu et al. (2022) would result in a stated preference experiment with too many attributes and overly complex choice tasks, which can increase the burden on respondents and reduce the quality of the collected data (Rose and Bliemer, 2009; Bliemer and Rose, 2024). Therefore, a smaller set of attributes is selected based on their relevance to the research objective and their suitability for the experimental design.

The selection follows two main criteria. First, the attributes should be supported by the pedestrian route choice and walkability literature discussed in Chapter 2. Second, they must be included in a way that avoids inconsistencies between textual or numerical information and the visual information shown in the images. For example, if an image shows an empty street while the description indicates high crowdedness, the respondent receives conflicting information. Such inconsistencies can make the choice task harder to interpret and introduce noise into the collected data. Table 3.1 provides an overview of the selected attributes and how they are represented in the experiment. The following sections explain the attribute selection in more detail.

Category	**Attribute**	**Type**	**Representation**
Trip Characteristics	travel time	Non-visual	Minutes
	Crossing type	Non-visual	Text
	Number of crossings	Non-visual	Count
	Waiting time	Non-visual	Minutes
Environment	Sidewalk	Visual	Image
	Greenery	Visual	Image
	Traffic	Visual	Image
	Crowdedness	Visual	Image
Socio-demographics	Age	Respondent	Survey
	Gender	Respondent	Survey
	Culture	Respondent	Survey
	Income	Respondent	Survey
	Mobility	Respondent	Survey
	Walking behaviour	Respondent	Survey
Context	Weather	Context	Fixed
	Time of day	Context	Fixed
	Travel company	Context	Fixed
	trip-purpose	Context	Scenario

Table 3.1: Overview of selected attributes relevant for pedestrian route choice

3.2.1. Trip Characteristics

Trip characteristics relate to the efficiency of taking a route. Distance and travel time are among the most commonly studied factors in pedestrian route choice, and pedestrians generally prefer shorter and faster routes (Basu et al., 2022; Prato, 2009). In this thesis, travel time is used as the main measure of route efficiency. This choice is made because travel time is easier for respondents to interpret in a survey setting. Distance requires respondents to translate metres into an expected walking effort or duration, which may differ between individuals. Travel time therefore provides a more direct and intuitive measure of route efficiency.

In addition to travel time, road crossings are considered an important attribute. The literature indicates that crossings can influence route choice by affecting both perceived delay and safety, because they involve waiting times and interactions with traffic (Tong and Bode, 2022; Duives, 2025; Basu et al., 2022). This attribute can be represented in different ways, such as by the type of crossing, the waiting time, or the number of crossings along a route. How these aspects can be represented in the experiment is further described in Section 3.4.

3.2.2. Built- and Natural-Environment Factors

Pedestrian route choice is influenced by environmental characteristics that affect perceptions of comfort, safety, and attractiveness (Tong and Bode, 2022; Basu et al., 2022; Borst et al., 2008). In this study, these attributes are mainly represented through street-level images, in which many of the attributes described by Basu et al. (2022) are visible. A smaller set of built- and natural-environment factors is used to support the selection of images. These factors are chosen based on their relevance for walkability in urban settings and their ability to represent different types of street environments. Characteristics such as aesthetics, cleanliness, openness, and lighting are identified as relevant in the literature, but are not included as separate attributes. This is because they are difficult to measure and identify separately in a consistent way. Instead, they are captured through the overall appearance and composition of the images.

Based on the literature, the selected image attributes consist of four main environmental factors: sidewalk availability, greenery, traffic exposure, and crowdedness. Sidewalk availability refers to the presence and quality of pedestrian infrastructure, which affects accessibility and comfort (Borst et al., 2008; Duives, 2025). Greenery includes visible vegetation and is generally associated with more attractive walking environments (Basu et al., 2022; Choi and Kang, 2025). Traffic exposure refers to the presence of motorised vehicles and is often linked to perceived safety (Basu et al., 2022). Crowdedness describes the number of people in the environment and the perceived density of the street. The effect of crowdedness on route choice is not always consistent in the literature, because it can influence both perceived safety and comfort depending on the context (Basu et al., 2022; Duives, 2025). In this thesis, crowdedness is included because it can be identified from street-level images and is relevant in urban environments, where pedestrian densities are usually higher.

Weather and time of day can influence how pedestrians perceive and evaluate their environment (Tong and Bode, 2022; Duives, 2025). In particular, time of day can affect perceived safety and route choice behaviour, because pedestrians may evaluate environments differently during the day and at night (Basu et al., 2023). However, in this thesis, these factors are not the main focus and are therefore kept constant across all choice tasks. Other sensory factors, such as noise and smell, are also relevant to pedestrian experience (Duives, 2025). However, they cannot be derived from images or represented consistently in this experiment and are therefore not included.

3.2.3. Socio-Demographic Factors

In addition to route attributes, personal characteristics can also influence how pedestrians evaluate routes (Basu et al., 2022; Duives, 2025). These characteristics are not included directly in the choice tasks. Instead, they are collected separately to describe the respondent group and to provide context for interpreting the results.

The selected characteristics are age, gender, cultural background, income, mobility level, and walking behaviour. Age and mobility level may affect walking ability and sensitivity to environmental conditions (Borst et al., 2008). Gender is often related to differences in perceived safety (Mazzulla et al., 2024). Spatial and cultural context may influence walking behaviour and sensitivity to built-environment characteristics (Levine and Norenzayan, 1999; Lin et al., 2015). Income is included as a general socio-economic factor, and walking behaviour reflects habits that may influence route choice (Ton et al., 2019). These characteristics are collected through general survey questions.

To keep the choice tasks clear and consistent, travel company is included as a fixed contextual factor. It refers to whether a pedestrian walks alone or with others, which can influence route choice due to social interactions (Basu et al., 2022). In this thesis, respondents are asked to imagine that they are walking alone.

3.2.4. Contextual Factors

Finally, trip purpose is included to represent different walking situations. Trip purpose is not a route attribute itself, but it can influence how route attributes are evaluated. Previous studies show that walking behaviour differs by trip purpose, for example in terms of walking distance, walking duration, and the role of the built environment (Larranaga and Cybis, 2014; Hatamzadeh et al., 2014; Yang and Diez Roux, 2012; Watson et al., 2021; Mondschein, 2018). Some trips, such as walking to public transport, work, or school, tend to be more time-constrained. Other trips, such as walking in free time, are more flexible and may allow more attention to the quality of the walking environment (Larranaga and Cybis, 2014; Agrawal et al., 2008). These differences suggest that route attributes may be evaluated differently depending on how goal-oriented or time-sensitive the walking trip is. The specific trip-purpose contexts considered in this thesis are defined in Section 3.4.

Area familiarity is also recognised as an important factor in the literature, because pedestrians rely on their knowledge of the environment and their habits when choosing a route (Duives, 2025). However, because the experiment uses random street-level images from multiple cities in the Netherlands, familiarity cannot be included in a meaningful way. It is therefore excluded from the experiment.

3.3. Street-Level Image Data Collection

To analyse the influence of visual street characteristics on pedestrian route choice, a dataset of street-level images is constructed. This dataset is used to represent the environmental attributes described in Section 3.2. The aim is to capture a wide range of urban walking environments in the Netherlands. Following earlier applications of the CV-DCM, this thesis aims to construct a dataset of approximately 10,000 street-level images (van Cranenburgh and Garrido-Valenzuela, 2025; Terra et al., 2025). There is no exact required number of images. However, a larger and more diverse dataset provides more variation in visual street characteristics and helps the model learn visual patterns across different types of urban walking environments.

The dataset is constructed using the ten largest cities in the Netherlands: Amsterdam, Rotterdam, Den Haag, Utrecht, Eindhoven, Groningen, Almere, Tilburg, Breda, and Nijmegen. These cities are selected because they represent denser urban areas, different street designs, and a range of pedestrian environments across the country. For collecting street-level imagery, both Mapillary and Google Street View were explored (Google, 2026; Mapillary, 2026). Mapillary provides user-generated images with broad spatial coverage, but the image quality and consistency were not sufficient for this study. Google Street View provides higher and more consistent image quality, as well as more reliable metadata. Therefore, Google Street View images are used as the main data source. However, the methods described in this section could also be applied to other street-level image sources.

The construction of the image dataset follows four main steps: extracting pedestrian networks, sampling street-level image locations, requesting and filtering Google Street View images, and categorising the images based on the selected environmental attributes. These steps are explained in further detail in the next sections.

3.3.1. Pedestrian Network Extraction

For the ten largest cities in the Netherlands, a pedestrian network is constructed using OpenStreetMap data through the OSMnx library (OpenStreetMap contributors, 2017). The networks are filtered to include only pedestrian-accessible street segments based on OpenStreetMap tags. In general, street types such as footways, pedestrian streets, and residential roads are included. Segments that are not accessible to pedestrians, such as highways, are excluded.

After filtering, the networks are simplified using the built-in simplification procedure in OSMnx. The resulting network consists of nodes and edges. Nodes represent intersections, while edges represent walkable street segments. The full network extraction procedure and the exact inclusion and exclusion criteria are described in Appendix B.1. Figure 3.2 shows the pedestrian networks for Amsterdam, Rotterdam, and Den Haag. Each city is assigned a central reference location, such as a main square, shown as a red dot. These networks form the basis for the image sampling process. The locations used as central reference points are also shown in Appendix B.1.

(a) Amsterdam **(b)** Rotterdam **(c)** Den Haag

Figure 3.2: Pedestrian networks for the three largest cities in the Netherlands

3.3.2. Street-Level Image Sampling Strategy

Street-level images are sampled from the pedestrian networks of the ten largest cities to represent the visual walking environment in an urban context. The sampling strategy is designed to capture variation in street-level characteristics that pedestrians are likely to experience. The formal sampling procedure is described in Appendix B.2.

To reflect the spatial structure of the pedestrian network, streets are sampled with a probability proportional to their length. This ensures that longer street segments are not underrepresented compared with shorter ones. In addition, a centre-weighted sampling approach is applied. Initial uniform sampling showed that denser street environments were underrepresented. Because these environments are relevant for this study, streets closer to the city centre are given a higher sampling probability. This is done using distance-based weighting, where the chance of a street being selected gradually decreases with its distance from the city centre. This increases the representation of central areas without excluding the wider urban network.

Images are sampled in small groups along the same street segment. Each group consists of three nearby locations with a consistent viewing direction aligned with the street. This allows one street environment to be represented by multiple observations. It also makes it possible to compare images within the same street segment.

Image availability and quality are checked using Google Street View metadata. Images are restricted to daytime conditions and to months with comparable seasonal characteristics. This reduces variation caused by factors such as snow or seasonal differences in vegetation. In this study, only images captured between May and September are used. Images with invalid panoramas, extreme tilt, or duplicate viewpoints are excluded. Using this approach, a total of 10,727 street-level images are sampled across the pedestrian networks. Figure 3.3 shows the spatial distribution of the sampled image locations for Amsterdam. A higher concentration of points is visible around the city centre, reflecting the centre-weighted sampling approach. Locations across the wider urban area are still included.

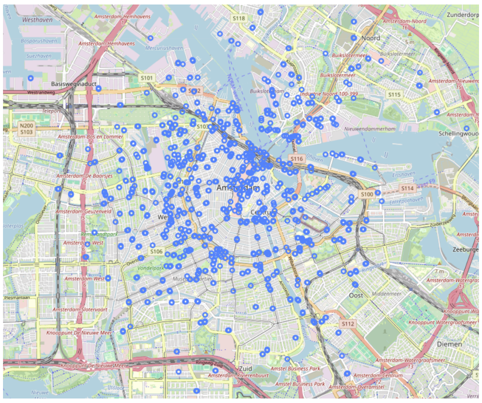


Figure 3.3: Spatial distribution of sampled Street View image locations in Amsterdam

3.3.3. Image Analysis and Filtering

After sampling the street-level images, the dataset is analysed and filtered to ensure that the images are suitable for representing pedestrian street environments in the SP experiment. In addition to official Google Street View images, the dataset includes user-contributed images. While this increases coverage of pedestrian streets, it also introduces images that are not appropriate for the experiment. The image analysis and filtering strategy is therefore designed to only keep images that represent walkable street environments, while removing images that could introduce noise or bias. The full pseudocode and threshold values are provided in Appendix B.3.

Extraction of Visual Features

Visual features are extracted from each image using two computer vision models. First, a semantic segmentation model is used to assign each pixel to a visual class. Mask2Former, trained on the Mapillary Vistas dataset, is used for this purpose (Cheng et al., 2022; Neuhold et al., 2017). This allows images to be represented using pixel shares for all segmentation classes. For the later filtering and modelling steps, only the classes that are relevant for identifying street environments and describing the pedestrian walking environment are used, including road, sidewalk, vegetation, sky, and water.

Second, an object detection model is used to identify and count pedestrians and motorised vehicles using YOLOv8 (Ultralytics, 2023). While the segmentation model can also identify these elements, it represents them only as pixel shares. Object detection provides more reliable counts of individual pedestrians and vehicles, which are used in later filtering and categorisation.

Filtering Based on Image Content

The extracted visual features are used to filter images through a set of rule-based criteria. Images with low brightness are removed to exclude night-time or low-visibility conditions. Although an initial brightness filter was applied during data collection, additional filtering was needed after inspecting the dataset, because some low-quality images remained. Images that do not represent street environments are also excluded. This includes indoor scenes, images dominated by rail infrastructure, and images mainly showing natural elements such as water or sky. These cases are identified using combinations of pixel shares derived from the segmentation results. Examples of accepted and rejected images are shown in Figure 3.4.

The remaining images are then evaluated for pedestrian accessibility based on the presence of sidewalks or pedestrian areas. Images that meet these criteria are classified as pedestrian-accessible streets, while all other images are assigned to the category `other`. The thresholds used for this classification were defined through iterative testing. A complete overview of the filtering rules and thresholds is provided in Appendix B.2, and an overview of some of the accepted and rejected images is included in Appendix C.1.

(a) Accepted image **(b)** Rejected image: indoor **(c)** Rejected image: nature

Figure 3.4: Examples of accepted and rejected street-level images

Final Validation

After applying the rule-based filtering, a small number of unusable images remained in the accepted image set. Therefore, a final manual check was performed, in which images that clearly did not represent pedestrian street environments were removed. This included, for example, indoor images or images without any visible road. This step was kept limited to preserve the reproducibility of the filtering process, while still improving the overall data quality. The final distribution of images across categories is shown in Table 3.2.

Category	**Count**
Pedestrian-accessible	6189
Other	4094
Rejected indoor	376
Rejected rail	43
Rejected light	15
Rejected nature	10

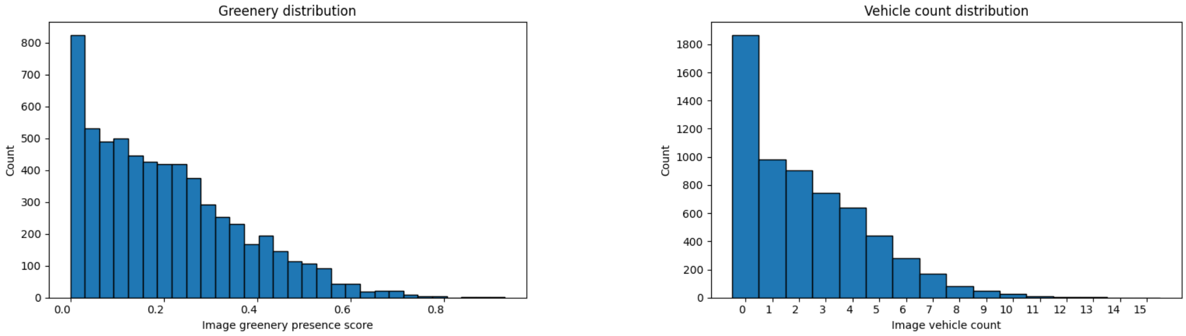
Table 3.2: Number of images per category after filtering

3.3.4. Construction of Visual Categories

The images are categorised to balance the images in the experiment and to create clear and interpretable visual differences. The categorisation supports the survey design, meaning that each category should contain enough images so that multiple distinct examples can be selected per category. The filtered images are categorised based on three visual characteristics: greenery presence, vehicle presence, and people presence. These characteristics are selected because they are clearly visible in images and can influence how pedestrians perceive street environments. Although sidewalk availability was identified as an important attribute in Section 3.2, it is not included as a separate visual category. This is because it was difficult to reliably identify consistent and visually distinct categories based on sidewalk presence using the image data.

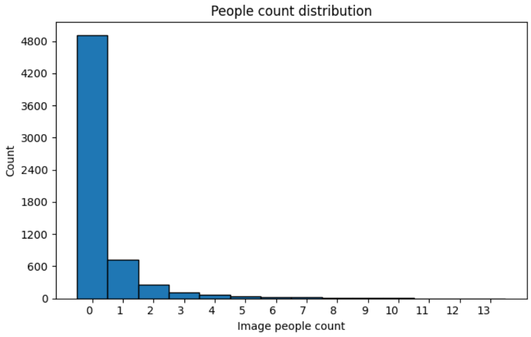
The construction of these categories follows the process described in Appendix B.4. Only images classified as pedestrian-accessible are used. For these images, image-level variables are computed based on the outputs of the image analysis step. These include the number of detected vehicles, the number of detected people, and a greenery presence score based on the share of vegetation and terrain in the image. The definitions of these variables are provided in Appendix B.3.

To determine suitable category definitions, the distributions of the variables are analysed, as shown in Figure 3.5. Several binning approaches were explored, including equal-width bins, quantile-based bins, manual thresholds, and clustering methods. The final approach is selected based on reproducibility, interpretability, whether the resulting categories contain enough images, and whether they are visually distinguishable. Since the distributions of vehicle and people presence are strongly skewed, a binary categorisation is used for these variables. This means that images are classified as either having zero or non-zero counts. For greenery presence, a quantile-based categorisation is applied, resulting in three groups: low, medium, and high greenery presence. The full binning rules are provided in Appendix B.3.



(a) Image greenery presence score

(b) Image vehicle presence count



(c) Image people presence count

Figure 3.5: Distributions of visual characteristics of filtered images

The resulting combinations of these categories are shown in Table 3.3. These combinations form the basis for the selection of images used in the survey. While some combinations occur more often than others, the categorisation ensures that all categories contain enough images to be represented, while still capturing a wide range of visual environments. Visual examples of the categories are provided in Appendix C.2.

Vehicle presence	Greenery presence	People presence	Count
Present	Medium	None	1297
Present	High	None	1220
Present	Low	None	1057
None	High	None	539
None	Low	None	410
None	Medium	None	383
Present	Low	Present	310
None	Low	Present	286
Present	Medium	Present	256
Present	High	Present	185
None	Medium	Present	127
None	High	Present	119

Table 3.3: Counts for combined image characteristics

3.4. Stated Preference Experiment

This part of the research addresses the second sub-question: *How can the trade-offs pedestrians make between visual and non-visual route attributes be measured?* It explains the Stated Preference (SP) approach, the survey design, the choice task structure, and the implementation of the final survey.

To analyse these trade-offs, pedestrian route choice preferences are collected using a SP experiment in the form of an online survey. Respondents are presented with hypothetical route choice situations in which they choose between two street alternatives. These alternatives differ in travel time and visual street-level characteristics. The visual characteristics are represented using the street-level image dataset described in Section 3.3.

The final survey design is developed through mock survey testing and a pilot survey. This is important because the choice tasks combine multiple types of information, including numerical attributes, trip-purpose contexts, and street-level images. The mock survey and pilot survey are used to evaluate whether respondents understand the choice tasks, interpret the attributes as intended, and are able to make meaningful trade-offs between alternatives (Bliemer and Rose, 2024). The mock survey designs are evaluated through talk-through sessions with peers and colleagues, while the pilot survey is used to test the complete online survey setup before finalising the design. The findings from these tests are used to refine the final survey design, which is described in the following sections.

3.4.1. Stated Preference Approach

In this thesis, an SP approach is used to collect data for pedestrian route choice models. This approach is chosen because it allows control over the choice situations shown to respondents. Since the aim is to analyse trade-offs between travel time and visual street-level characteristics, and to examine whether these trade-offs differ across trip-purpose contexts, both the route alternatives and the trip-purpose contexts need to be defined in advance and varied systematically.

As described in Section 2.1.3, both Revealed Preference (RP) and Stated Preference (SP) data can be used to estimate route choice models. RP data is based on observed behaviour and can show which routes pedestrians choose in real-world situations. However, for the purpose of this thesis, RP data is less suitable. With RP data, it is often unclear which routes were considered and whether the route was chosen because of travel time, visual quality, familiarity, habit, or other unobserved factors. In addition, detailed contextual information is often limited. For example, Sevtsuk et al. (2021) note that their GPS-based data did not include identifiable user attributes or information about trip origin type, destination type, or trip purpose. This prevented them from analysing differences in pedestrian route choice preferences by trip purpose. Since trip purpose is central in this thesis, this is an important limitation of using RP data.

For these reasons, a SP experiment is used to collect pedestrian route choice preference data. Since the choice task is framed as two street alternatives within the same walking context, it is known which options the respondent considers. However, a limitation of this approach is that the choices are made in hypothetical situations and may differ from actual walking behaviour (Ben-Akiva and Lerman, 1985; Brownstone et al., 2000). Since the goal of this thesis is to analyse stated route choice preferences, and not to directly predict actual route choice behaviour, this limitation is accepted.

The SP experiment is implemented as an online survey using PixelSurvey (Garrido-Valenzuela et al., 2025). This software is used to design a web-based survey in which street-level images can be included directly within the choice tasks. The online format also supports data collection, because the survey can be distributed more easily. This is relevant for the CV-DCM, because modelling preferences related to visual street-level characteristics requires many choice observations across different images (van Cranenburgh and Garrido-Valenzuela, 2025).

3.4.2. Survey Design

The survey is structured so that respondents complete the choice tasks early on, because these tasks form the main part of the data collection. Placing the choice tasks near the beginning helps reduce the effect of respondent fatigue, which can affect response quality when more demanding tasks are placed later in the survey (Bliemer and Rose, 2024). The structure of the survey is as follows:

1. **Introduction and informed consent**
   The survey starts with an introduction explaining the purpose of the study, followed by an informed consent statement. This ensures that respondents understand the study and agree to participate.
2. **Demographic and screening questions**
   Respondents are asked a small number of general questions, such as their age, gender, whether they live in the Netherlands, and whether they are able to walk independently. These questions are used to check whether respondents meet the requirements of the study and to describe the sample.
3. **Choice tasks**
   This is the main part of the survey and consists of 12 route choice tasks. In each task, respondents choose between two street alternatives that differ in travel time and visual street-level characteristics. The choice tasks are divided across trip-purpose contexts, which are described in Section 3.4.3.
4. **Perception questions**
   After the choice tasks, respondents are asked to rate individual street images based on their first impression, using a Likert scale from 1 to 5. These questions are included to capture how different street environments are perceived, without comparing them directly to other alternatives.
5. **Walking behaviour and background questions**
   Lastly, respondents answer questions about their walking habits and background. These responses are used to describe the participant group and can be used to analyse possible differences in preferences between respondents.

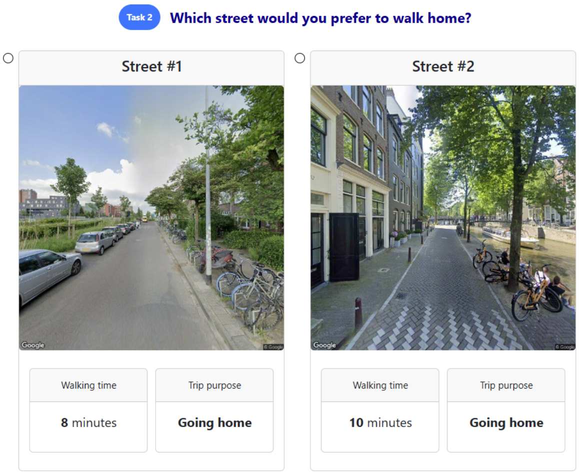
3.4.3. Choice Task Design

The final survey uses choice tasks in which respondents choose between two street alternatives. Each alternative is defined by a travel time and a street-level image. The trip-purpose context is shown together with the attributes of the choice task, so that the respondent can clearly see the walking context in which the choice takes place. The final attributes are shown in Table 3.4. An overview of all questions in the final survey can be found in Appendix E.

Category	Attribute	Levels	Representation
Trip Characteristics	Travel time	6, 8, 10	Minutes
Environment	Vehicles People Greenery	None, present None, present Low, medium, high	Image
Context	trip-purpose	Walking to public transport Walking to work or school Walking home Walking in free time	Text

Table 3.4: Attributes included in the final survey

The final choice task design is based on the selected route attributes in Section 3.2, the image categories developed in Section 3.3.4, and the results of the mock and pilot survey. The design aims to include attributes that are relevant for pedestrian route choice, can be clearly interpreted by respondents, and can be represented consistently in the choice tasks. An example of a choice situation is shown in Figure 3.6.



Task 2 Which street would you prefer to walk home?

○ Street #1

Walking time: 8 minutes

Trip purpose: Going home

○ Street #2

Walking time: 10 minutes

Trip purpose: Going home

Figure 3.6: Example of a choice task in the context of walking home.

Trip Characteristics and Environment

Travel time is included as the only trip characteristic in the final choice tasks. As described in Section 3.2, travel time is one of the main measures of route efficiency and is commonly used in route choice studies. The final survey uses travel time levels of 6, 8, and 10 minutes. These levels create variation in route efficiency, while avoiding very large differences between alternatives. This is supported by the pilot survey, where the differences between 5, 8 and 11 minutes were considered too large. This could lead to alternatives being chosen mainly because one alternative was much faster.

The environmental characteristics are represented through street-level images. The images use the categories described in Section 3.3.4: vehicle presence, people presence, and greenery level. These characteristics are used because they are clearly visible in the image data and relate to environmental factors discussed in the pedestrian route choice literature, such as traffic conditions, crowdedness, and natural elements (Basu et al., 2022; Duives, 2025). Respondents are only shown the image itself and are not informed about the underlying image categories.

Crossing-related attributes are not included in the final choice tasks, even though crossings are considered relevant for pedestrian route choice because they can influence delay, safety, and route attractiveness (Basu et al., 2022; Tong and Bode, 2022; Duives, 2025). The mock and pilot survey showed that crossing attributes were difficult to interpret in combination with the street-level images. In some cases, the crossing information conflicted with what was visible in the image. For example, a zebra crossing could be visible in the street-level image, while the crossing attribute indicated an unmarked crossing. This led to confusion about which information respondents should rely on and reduced the clarity of the choice tasks.

Trip-purpose Contexts

The final survey uses four trip-purpose contexts that reflect the application of the study, the trip-purpose literature, and the pilot survey results. These contexts are included because pedestrian route choice preferences are expected to differ depending on how time-sensitive and goal-oriented a walking trip is. This may influence how pedestrians prioritise different route attributes.

Walking to public transport is included as the most time-sensitive and goal-oriented context. In this situation, pedestrians often need to arrive on time, meaning that travel time and distance are expected to be important, because missing a connection can affect the rest of the trip. However, pedestrians walking to public transport stations may still consider safety, sidewalk availability, and aesthetic qualities (Agrawal et al., 2008).

Walking to work or school is also goal-oriented, but the time pressure can be considered less strict than when catching public transport. Travel time and distance remain important, but pedestrians may be more willing to consider visual street-level characteristics (Larranaga and Cybis, 2014; Hatamzadeh et al., 2014).

Walking home is included as a separate context because it represents a common return trip. It is still destination-oriented, but is expected to be less time-sensitive than walking to public transport or to work or school. This means that pedestrians may have more room to consider visual street-level characteristics when choosing a route (Watson et al., 2021).

Finally, walking in free time is included as the most flexible context. It represents leisure or recreational walking, where the destination and arrival time are usually less fixed. Walking is part of the experience, not just a mode of transport. In this context, visual street-level characteristics, such as the overall quality and attractiveness of the walking environment, are expected to play a larger role in route choice preferences (Yang and Diez Roux, 2012; Larranaga and Cybis, 2014).

Priors and Experimental Design

The final choice tasks are generated using an efficient experimental design. Designing these tasks requires selecting combinations of attribute levels that make it possible to estimate trade-offs between attributes. Because each respondent can only be given a limited number of choice tasks in the survey, it is important to obtain as much information as possible from each observation. For this reason, an efficient experimental design is used, as this type of design is specifically developed to improve the statistical efficiency of parameter estimation in stated choice experiments (Rose and Bliemer, 2009).

The efficient design is generated using Ngene. This software searches for combinations of attribute levels that minimise the D-error of the design. The D-error represents the expected uncertainty in the parameter estimates. A lower D-error indicates that the design is expected to produce more precise estimates for the assumed model and priors (Rose and Bliemer, 2009). The design gives a specification for each choice task alternative based on the final attributes presented in Table 3.4. The full Ngene specifications for the final experimental design are provided in Appendix E.

The priors used for the final experimental design are based on the reviewed literature and the pilot survey. The initial priors for the pilot design followed the literature discussed in Section 2.2. Longer travel times, vehicle presence, people presence, and lower levels of greenery were expected to reduce utility. The value used for the initial travel time prior was also based on a study on the value of travel time for walking in the Netherlands (Kouwenhoven et al., 2023). The pilot survey estimates followed these expected directions and showed early differences between trip-purpose contexts, as shown in Appendix D.

Because the pilot sample size was small, with only 16 respondents, the estimated parameter values are considered uncertain. Using strongly informative priors in this situation could bias the experimental design (Bliemer and Collins, 2016). Therefore, the pilot-based priors are not used directly. Instead, they are scaled down to keep the expected signs, while limiting their influence on the final design. An additive coding approach is used for the combined image categories, consistent with the pilot design. The adjusted priors are used only to generate an efficient experimental design and should not be interpreted as final results. The final prior values are reported in Table 3.5, and the complete Ngene code is included in Appendix E.

Parameter	Pilot estimate	Final prior
Travel time	-0.278	-0.15
Vehicle present	-0.450	-0.25
Greenery low	-0.543	-0.30
Pedestrian present	-0.501	-0.25

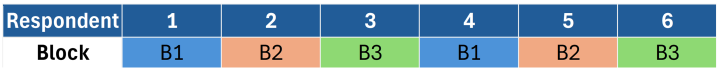
Table 3.5: Pilot estimates and adjusted priors for the final design

The final experimental design was selected based on the lowest D-error achieved after running the optimisation in Ngene for 30 minutes. Because several parameters were changed after the pilot survey, the first design, with 24 rows divided over 2 blocks, resulted in a relatively high D-error of 1.02. Therefore, a second design, with 36 rows divided over 3 blocks, was tested. This resulted in a lower D-error of 0.65. Since this design was more efficient and also provided more variation in the choice tasks, the 36-row design was selected.

3.4.4. Survey Implementation

In the final survey, choice tasks, trip-purpose contexts, and images are assigned to respondents using a structured method. This is important because the same types of choice tasks should not always appear in the same trip-purpose context, as this could systematically affect the results. Therefore, the final survey varies how tasks and contexts are combined across respondents.

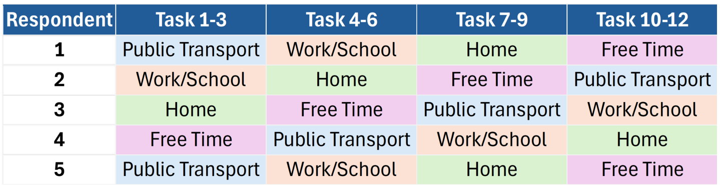
Each choice task represents one choice between two street alternatives. Each alternative is described by a travel time and a street-level image. The final Ngene design consists of three blocks, with 12 choice tasks in each block. Respondents are assigned to one of the three blocks in a repeating sequence, as shown in Figure 3.7.



Respondent	1	2	3	4	5	6
Block	B1	B2	B3	B1	B2	B3

Figure 3.7: Cyclic assignment of respondents to design blocks

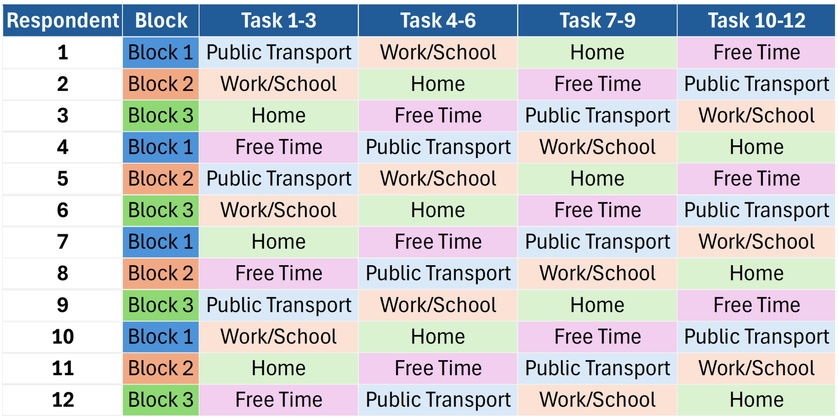
Within each block, the 12 choice tasks are divided into four sets of three tasks. Each set is assigned to one of the four trip-purpose contexts: walking to public transport, walking to work or school, walking home, and walking in free time. To avoid that the same task set is always linked to the same context, this assignment is shifted across respondents. This rotation is shown in Figure 3.8. As a result, the same group of choice tasks is evaluated under different trip-purpose contexts by different respondents.



Respondent	Task 1-3	Task 4-6	Task 7-9	Task 10-12
1	Public Transport	Work/School	Home	Free Time
2	Work/School	Home	Free Time	Public Transport
3	Home	Free Time	Public Transport	Work/School
4	Free Time	Public Transport	Work/School	Home
5	Public Transport	Work/School	Home	Free Time

Figure 3.8: Rotation of choice task sets across trip-purpose contexts

The block assignment and context rotation operate independently. The block cycle has a length of three, while the context rotation has a length of four. This means that the full assignment pattern repeats after 12 respondents, and that each design block is combined with each context rotation once. This reduces systematic bias between the task content and the trip-purpose context. The combined structure is shown in Figure 3.9.



Respondent	Block	Task 1-3	Task 4-6	Task 7-9	Task 10-12
1	Block 1	Public Transport	Work/School	Home	Free Time
2	Block 2	Work/School	Home	Free Time	Public Transport
3	Block 3	Home	Free Time	Public Transport	Work/School
4	Block 1	Free Time	Public Transport	Work/School	Home
5	Block 2	Public Transport	Work/School	Home	Free Time
6	Block 3	Work/School	Home	Free Time	Public Transport
7	Block 1	Home	Free Time	Public Transport	Work/School
8	Block 2	Free Time	Public Transport	Work/School	Home
9	Block 3	Public Transport	Work/School	Home	Free Time
10	Block 1	Work/School	Home	Free Time	Public Transport
11	Block 2	Home	Free Time	Public Transport	Work/School
12	Block 3	Free Time	Public Transport	Work/School	Home

Figure 3.9: Combined block assignment and context rotation across respondents

The order of the choice tasks within each block follows the Ngene design and is kept fixed in the survey implementation. The rotation only affects the link between task sets and trip-purpose contexts, not the order in which the choice tasks are shown.

Image Assignment

After the choice task design is created, images are assigned to the choice tasks based on the image categories defined in Section 3.3.4. Each image category represents a combination of vehicle presence, people presence, and greenery level. The Ngene design only specifies which image category should be shown for each alternative, not the exact image. Therefore, a separate image assignment step is needed to link specific images to the categories in the design.

The image assignment limits image repetition as much as possible. Images are selected from the corresponding image category, and each image is used at least once before repetition occurs. As a result, respondents do not see the same image twice within the choice tasks. This is important because repeated images could lead respondents to recognise earlier tasks or compare their answers across tasks. The position of the alternatives is also randomised, so that images are not always shown on the same side. This helps reduce positional bias, where respondents may prefer the left or right option regardless of the attributes shown (Bliemer and Rose, 2024).

3.5. Modelling Framework

This section addresses the third sub-question of this thesis: *How can visual and non-visual attributes be incorporated into a discrete choice model for pedestrian route choice preferences?* It introduces the model specifications and explains how the models are estimated and optimised.

To answer this question, the SP data is analysed using three model specifications: a baseline Multinomial Logit (MNL) model, a pixel-share MNL model, and a Computer Vision-Enriched Discrete Choice Model (CV-DCM). Other model structures were also considered to account for differences between respondents. As shown in the literature, pedestrian route choice preferences may differ by age, gender, walking ability, or other individual characteristics (Basu et al., 2022; Borst et al., 2008; Mazzulla et al., 2024). Models such as mixed logit or latent class models can capture these differences (Train, 2009). However, the aim of this thesis is to compare different ways of including visual street-level information in the utility function and to examine how route choice preferences differ across trip-purpose contexts. Including individual-level heterogeneity would add another modelling dimension. These differences are therefore acknowledged, but not explicitly modelled in this thesis.

This section first introduces the discrete choice modelling framework and the utility formulation. Next, the three model specifications are described, including the attributes used in the utility function. Finally, the model estimation and optimisation procedures are explained. For the pixel-share MNL, this includes alternative model specifications. For the CV-DCM, this includes the optimisation and hyperparameter tuning process.

3.5.1. Discrete Choice Modelling Framework

DCMs are commonly used in route choice research because they provide a clear and theoretically grounded way to represent how individuals choose between alternatives based on the attributes included in the utility function (Ben-Akiva and Lerman, 1985; Prato, 2009). Since research shows that pedestrians sometimes accept a longer route when another route is perceived as more attractive, this framework is used in this thesis to represent trade-offs between travel time and perceived visual street-level characteristics (Sevtsuk et al., 2021; López-Lambas et al., 2021).

At the core of this modelling approach is Random Utility Theory (RUT). Within this theory, choices are represented as the outcome of Random Utility Maximisation (RUM). In this thesis, this means that pedestrians are assumed to choose the route alternative that gives them the highest utility. This assumption aligns with behavioural theories of pedestrian route choice, which describe route choice as a decision based on perceived costs and benefits rather than only on objective measures such as distance (Hoogendoorn and Bovy, 2004; Duives, 2025). In discrete choice modelling, it is assumed that not all factors influencing a choice can be observed or measured. Utility is therefore represented as the sum of an observable component V_{in} and an unobservable component ε_{in}:

$$U_{in} = V_{in} + \varepsilon_{in} \tag{3.1}$$

The observable part of utility is based on variables explicitly included in the model, while the unobservable part captures influences that are not included in the data. The systematic part of utility is specified as a function of the observed attributes:

$$V_{in} = \sum_{m} \beta_m \cdot x_{imn} \tag{3.2}$$

Here, x_{imn} represents attribute m of alternative i for individual n, such as travel time or the amount of greenery in an image. The parameter β_m indicates how strongly attribute m affects utility. This reflects how individuals trade off different route characteristics (Ben-Akiva and Lerman, 1985; Train, 2009). For example, a negative parameter for travel time indicates that longer travel times reduce utility.

Because part of utility is unobservable, choices cannot be predicted exactly. Instead, this framework assumes that alternatives with higher utility are more likely to be chosen (McFadden, 1974). To translate utility into choice probabilities, all models used in this thesis assume that the unobserved error terms are independently and identically distributed according to an extreme value type I distribution with variance $\frac{\pi^2}{6}$ (Ben-Akiva and Lerman, 1985). Under this assumption, the probability that individual n chooses alternative i from choice set C_n is given by:

$$P_{in} = \frac{e^{V_{in}}}{\sum_{j \in C_n} e^{V_{jn}}} \tag{3.3}$$

This means that the probability of choosing a route alternative depends on how its utility compares with the utilities of the other available alternatives. Alternatives with higher utility have a higher probability of being chosen. This probability formulation is used for all three model specifications. However, the way V_{in} is calculated differs per model.

3.5.2. Model 1: Baseline MNL

The first model is a simple MNL model that only uses travel time as an explanatory variable. Travel time is included because it is one of the most commonly used attributes in pedestrian route choice modelling and reflects the logic of a shortest-path approach (Basu et al., 2022; Prato, 2009). This model is used to examine whether the preferences observed in the experiment can be explained by travel time alone. It also serves as a simple reference model for evaluating the added value of including visual street-level characteristics.

3.5.3. Model 2: Pixel-Share MNL

The pixel-share MNL model extends the baseline MNL by including pixel-share variables as attributes in the utility function. The included variables are sidewalk share, pedestrian area share, vehicle share, vegetation share, and people share. These variables are selected based on the built- and natural-environment attributes identified in Section 3.2.2.

Sidewalk share and pedestrian area share are included because pedestrian infrastructure is expected to influence accessibility and comfort (Borst et al., 2008; Duives, 2025). Although both variables represent pedestrian space, they are separate and non-overlapping classes in the Mask2Former output. Vegetation share is included to represent greenery, which is generally associated with more attractive walking environments (Basu et al., 2022; Choi and Kang, 2025). In this model, vegetation share consists of both vegetation and terrain pixels, because terrain also includes elements such as grass fields. Vehicle share is included as a measure of traffic exposure, which is often linked to perceived safety (Basu et al., 2022). People share is included as a measure of crowdedness, because the number of people visible in the street environment can influence both comfort and perceived safety depending on the context (Basu et al., 2022; Duives, 2025).

In this pixel-share model, visual street-level characteristics are included explicitly. The model uses interpretable variables that directly reflect image characteristics. This makes it possible to analyse whether specific visual elements add explanatory value to the preferences observed in the choice experiment, and whether their effects differ across trip-purpose contexts.

3.5.4. Model 3: Computer-Vision Enriched Discrete Choice Model

The CV-DCM uses the full street-level image as input, allowing the model to learn which visual information is relevant for explaining the observed route choice preferences. This thesis follows the CV-DCM framework proposed by van Cranenburgh and Garrido-Valenzuela (2025), in which utility depends on both numerical attributes and information extracted from images. In this formulation, the observable part of utility consists of a component derived from numerical attributes, such as travel time, and a component derived from the street-level image:

$$U_{in} = \overbrace{\underbrace{\sum_{m} \beta_m \cdot x_{imn}}_{\text{Systematic utility derived from numeric attributes}} + \underbrace{\sum_{k} \beta_k \cdot z_{ikn}}_{\text{Systematic utility derived from attributes encoded in the image}}}^{V_{in}} + \epsilon_{in} \tag{3.4}$$

Here, x_{imn} represents the numeric attributes of alternative i for individual n, while z_{ikn} represents attributes derived from the street-level image. These image-based attributes are obtained through feature extraction:

$$Z_{in} = \phi(I_{in} \mid w) \tag{3.5}$$

The resulting feature vector Z_{in} does not have a direct interpretation, because it consists of a high-dimensional numerical representation of the full image. Rather than describing specific visual elements, it captures broader visual patterns in the street environment. These learned image-based attributes are included in the utility function and are estimated together with the behavioural parameters. The general structure of the CV-DCM is shown in Figure 3.10.

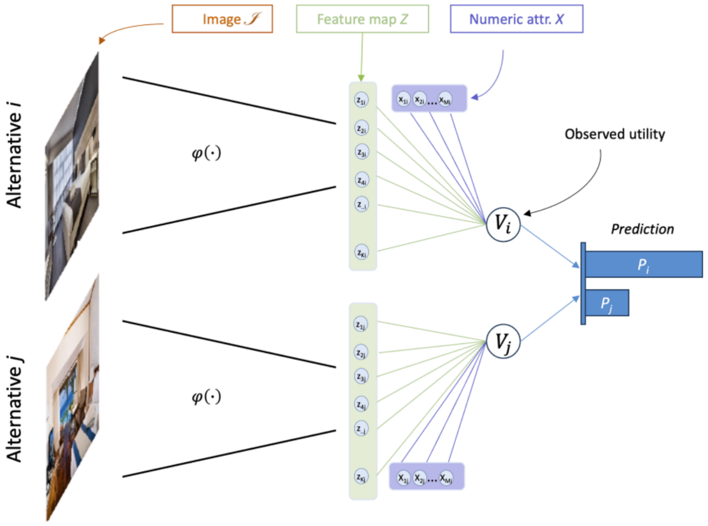


Figure 3.10: Model structure of the CV-DCM, as presented by van Cranenburgh and Garrido-Valenzuela (2025)

By using both the pixel-share MNL and the CV-DCM, two ways of including visual information can be compared. The pixel-share model uses visual variables that are defined in advance, while the CV-DCM learns a representation directly from the full street-level image. This makes it possible to examine whether predefined visual variables are sufficient to explain pedestrian route choice preferences, or whether the full image contains additional information.

3.5.5. Model Estimation and Optimisation

All models are implemented in Python. The baseline MNL model and the pixel-share MNL model are estimated using Biogeme. Biogeme is commonly used for estimating DCMs and provides standard tools for likelihood-based estimation and statistical output. The CV-DCM is implemented in PyTorch, following the framework of van Cranenburgh and Garrido-Valenzuela (2025).

The data is divided into a training set and a test set before model estimation. The training set is used to estimate the model parameters, while the test set is used to evaluate how well the models perform on unseen observations. The same train-test split is used for all model specifications, so that differences in predictive performance can be evaluated using the same test data. The models are estimated by fitting the model parameters to the collected SP data, so that the observed choices in the data are made as likely as possible. This is known as maximising the likelihood of the observed choices (Ben-Akiva and Lerman, 1985; Train, 2009). For each model type, five model instances are estimated. First, one combined model is estimated using all choice tasks. This provides an estimate of pedestrian route choice preferences aggregated across all trip purposes. In addition, four trip-purpose-specific models are estimated, where each model only uses observations from one trip-purpose context. Estimating separate models makes it possible to analyse whether the relative importance of travel time and visual street-level characteristics differs across contexts.

During model estimation, several optimisation steps are performed. This optimisation differs per model type. For the baseline MNL model, no further optimisation is applied, because this model only includes travel time and is used as a benchmark. For the pixel-share MNL model, different specifications are tested to check whether all visual variables should be included. For the CV-DCM, hyperparameter tuning is performed, because the training process depends on several settings that need to be selected. The evaluation measures described in Section 3.6 are also used during model optimisation and hyperparameter tuning, because they show how well the different model specifications explain the observed choices.

Train-Test Split

In this thesis, an image-level train-test split is used for model estimation and evaluation. In this approach, each image is assigned either to the training set or to the test set. The training set is used to estimate the model parameters, while the test set is used to evaluate predictive performance on unseen observations. The same train-test split is used for all model specifications, so that differences in predictive performance between the baseline MNL, pixel-share MNL, and CV-DCM can be evaluated using the same test data.

The train-test split should be created carefully when the data contains multiple observations from the same respondents and street-level images can be repeated across respondents. A respondent-level split assigns all observations from one respondent either to the training set or to the test set. This avoids respondent leakage and evaluates whether the model generalises to new respondents. This is a common strategy in stated choice data, where respondents complete multiple choice tasks (Bliemer and Rose, 2005). However, because the same street-level images can be shown to different respondents, a respondent-level split can result in the same images appearing in both the training and test set.

This image leakage is especially problematic for the CV-DCM. If the same image appears in both sets, the model could recognise the image. This means that the model may learn which specific image is preferred, rather than why the image is preferred. Therefore, an image-level split is used to avoid image leakage and to evaluate whether the model generalises to unseen route images. This follows the approach of van Cranenburgh and Garrido-Valenzuela (2025), who use an image-based split to avoid image leakage. The final split proportions and the resulting number of observations in the training and test sets are reported in Section 4.4.

Pixel-Share MNL Specification Checks

For the pixel-share MNL model, several additional specifications are estimated. The included pixel-share attributes are described in Section 3.5.3, but their relevance may differ between trip-purpose contexts. This means that some visual attributes may add explanatory value in one context, while adding little information or noise in another. Therefore, different combinations of pixel-share variables are tested to examine how this affects parameter significance and model performance.

In addition, a specification with an alternative-specific constant is estimated. This is done to test whether there is a general preference for one of the two alternatives that is not explained by the included attributes (Ben-Akiva and Lerman, 1985). The additional specifications are compared using the performance measures described in Section 3.6.

CV-DCM Hyperparameter Tuning

The CV-DCM is estimated differently from the other two models, because it uses the full street-level image as input. Following van Cranenburgh and Garrido-Valenzuela (2025), the DeiT base model is used as the feature extractor. The model is applied through transfer learning. This means that it starts from a version of DeiT that has already been pre-trained on ImageNet (Deng et al., 2009). This gives the model a starting point for recognising general visual patterns.

During estimation, the CV-DCM optimises both the neural-network weights w and the preference parameters β. As with the other models, the aim is to make the predicted choices match the observed choices as well as possible. This is done by minimising the cross-entropy loss, which is equivalent to maximising the log-likelihood. An L2 regularisation term is added to reduce the risk of overfitting by penalising large neural-network weights. The strength of this penalty is controlled by γ, as shown in Equation 3.6.

$$\hat{w}, \hat{\beta} = \arg\min_{w,\beta} \left(\underbrace{-\frac{1}{N} \sum_{n=1}^{N} \sum_{j=1}^{J} y_{nj} \log\left(P_{nj} \mid X_{nj}, I_{nj}, \beta, w\right) +}_{\text{Cross-Entropy Loss}} \quad \underbrace{\gamma \sum_{r} w_r^2}_{\text{L2 Regularisation}} \right) \tag{3.6}$$

Since computational resources were available through TNO, a more extensive tuning strategy could be used. Although the main aim of this thesis is not to develop a model that perfectly explains pedestrian preferences, the choice of hyperparameters can influence model performance, since these settings can affect how the neural network learns from the images. The tuned hyperparameters are the optimisation algorithm, learning rate, batch size, and weight decay.

Initial tuning values are based on earlier applications of the CV-DCM (van Cranenburgh and Garrido-Valenzuela, 2025; Terra et al., 2025). A broad search using the combined model is first used to explore a wider range of possible hyperparameter values. From this broad search, a more focused search is performed around the best-performing options. Because each trip-purpose model only uses part of the full sample, a final search is carried out separately for each trip-purpose model.

3.6. Model Evaluation

To analyse the effect of adding visual information, the models are evaluated using several methods. The baseline MNL model and the pixel-share MNL model are relatively interpretable, because their parameters are linked to clearly defined attributes. The CV-DCM is less directly interpretable, because it learns visual information from the full street-level image. Therefore, different evaluation methods are needed to assess whether the model outcomes are logical, whether they align with expectations, and whether unusual patterns occur. This is especially relevant for the trip-purpose-specific models, because the evaluation can show whether preferred street environments differ between walking contexts.

The models are evaluated in three steps. First, the behavioural consistency of the parameter estimates and the relative importance of the included attributes are evaluated. Second, the predictive performance of the models is compared using quantitative measures commonly used in discrete choice modelling (Ben-Akiva and Lerman, 1985; Train, 2009). These measures are used to assess how well the models explain the observed choices and whether adding visual information improves predictive performance. Finally, the results are examined through qualitative validation and a spatial application. The qualitative validation is used to assess whether the learned image utilities are visually and behaviourally meaningful. The spatial application then shows how the trip-purpose-specific models can be applied to compare preferred routes across a pedestrian network.

3.6.1. Parameter Estimates and Relative Importance

The first part of the evaluation focuses on the estimated parameters of each model. The sign of a parameter shows whether an attribute is associated with higher or lower utility. For example, a negative parameter for travel time indicates that longer routes are associated with lower utility. For the baseline MNL model and the pixel-share MNL model, the statistical significance of the parameters is also considered. Parameters with weak statistical support should be interpreted with caution, because they may not reflect consistent behavioural effects.

The relative importance of the attributes is also calculated. This is done to compare how much each attribute can influence utility within the same model. This is useful because the attributes are measured in different units, such as travel time in minutes and greenery as a pixel share. For each attribute m, the observed range is calculated using the lowest and highest values, denoted by $x_{m,\min}$ and $x_{m,\max}$. This range is multiplied by the estimated coefficient β_m. Since some attributes have a positive effect and others have a negative effect, the absolute value is used:

$$|\Delta U_m| = |\beta_m \cdot (x_{m,\max} - x_{m,\min})| \tag{3.7}$$

The relative importance of attribute m is then calculated by comparing its utility change to the total utility change of all attributes j included in the same model:

$$RI_m = \frac{|\Delta U_m|}{\sum_j |\Delta U_j|} \cdot 100\% \tag{3.8}$$

The result is expressed as a percentage. These percentages do not show how much each parameter contributes to total utility. Instead, they show how much each attribute contributes to the maximum possible utility difference within the model. This makes it possible to compare attributes based on their relative contribution within the same trip-purpose model. Because the calculation depends on the observed attribute ranges, the values should be interpreted as relative contributions within their specific dataset and model specification.

3.6.2. Predictive Performance Metrics

Model performance is evaluated on the test set using several quantitative measures to assess model fit, prediction accuracy, and model complexity (Ben-Akiva and Lerman, 1985). The hit-rate is used as a simple measure of prediction accuracy. It is the share of observations for which the model predicts the same alternative as the one that was chosen. However, the hit-rate only shows whether the predicted choice is correct. It does not show how certain the model was about the prediction.

The log-likelihood and cross-entropy are used as the main measures of predictive performance, because they are directly linked to the objective used during model estimation. The log-likelihood indicates how much probability the model assigns to the alternatives that were actually chosen. Higher log-likelihood values, meaning values closer to zero, indicate a better fit. Cross-entropy is the negative log-likelihood averaged over the number of observations. A lower cross-entropy therefore indicates better predictive performance. The Bayesian Information Criterion (BIC) is used to compare model fit while accounting for model complexity, where lower values indicate a better balance between fit and complexity (Schwarz, 1978). Finally, ρ^2 is used to compare the model with a null model without explanatory variables. A higher ρ^2 indicates a larger improvement over the null model.

3.6.3. Qualitative Model Validation

After the quantitative evaluation, the trip-purpose-specific models of the best-performing image-based model are analysed in more detail. Utility values depend on the scale and specification of each model, so the utility values cannot be compared directly between models (Ben-Akiva and Lerman, 1985; Train, 2009). Instead, the analysis focuses on the relative ranking of images within each trip-purpose model. This makes it possible to visualise and compare the street environments that receive higher or lower predicted image utility in each walking context.

Comparison with Perception Scores

First, predicted image utilities are compared with the perception scores collected in the survey. This is done to assess whether images with higher predicted utility are also perceived more positively by respondents. Because pedestrian route choice is influenced by how pedestrians perceive the walking environment, this comparison helps to assess whether the learned image utility reflects perceptions that are relevant for walking. However, the perception scores and predicted utilities are not the same. The perception scores describe how respondents rated individual images, while the predicted utilities are estimated from route choices. This comparison should therefore be seen as more of a consistency check.

Comparison of High- and Low-Utility Images

Images with high and low predicted utilities are compared to assess whether the model outcomes align with behavioural expectations. This also shows whether the highest- and lowest-utility images differ between trip-purpose contexts. For example, images with high predicted utility are expected to show more attractive or comfortable walking environments, while images with low predicted utility are expected to show less attractive walking environments.

Same-Street Consistency Check

The image collection process described in Section 3.3 aimed to collect three images per selected street segment. This means that multiple images are available from the same street with similar perspectives. These images are used for a same-street consistency check. This check shows whether the model gives similar utilities to similar street environments.

Image Augmentation Test

An exploratory image augmentation test is used to assess whether the image-based model responds to simple visual changes in a behaviourally expected way. Using generative AI, small changes are made to multiple images. The aim is to examine whether the predicted utilities change, whether these changes are consistent across images, and whether the changes differ between trip-purpose models.

The first two tests focus on greenery and traffic. These are important route choice attributes identified in Section 3.2, and they can be edited in an image using generative AI. Adding greenery is expected to increase image utility, while adding traffic is expected to decrease image utility (Basu et al., 2022; Duives, 2025). Two additional tests are also included. One test adds visible trash to the image. This is used to assess whether the model responds to a visual element that is associated with lower attractiveness, but is not explicitly included as a separate variable in the model. The other test adds pedestrian amenities in the form of benches. Benches are used as a simple test to assess whether the model responds to a small visual element related to pedestrian comfort.

3.6.4. Spatial Application

After the qualitative validation, the image-based model is applied in a spatial case study. As discussed in Section 2.2, pedestrians may deviate from the shortest route when other route characteristics are taken into account (Prashker and Bekhor, 2004; Prato, 2009; López-Lambas et al., 2021; Sevtsuk et al., 2021). The aim of this application is not to predict pedestrian flows. Instead, it shows whether there are differences between the shortest route and the routes preferred by the models, and how visual street-level characteristics can influence the route preferences.

Case Study Area

For the application, a neighbourhood in Amsterdam-Zuid is selected as the case study area. This area is chosen because it includes different types of streets, such as parks, residential streets, university surroundings, hospital areas, station environments, and more car-oriented streets. This variation makes it suitable for examining whether the model assigns different utilities to different urban environments. The area is also selected because TNO expressed interest in this location.

Pedestrian Network and Street-Level Images

The pedestrian network and street-level images are collected using an approach similar to the one described in Section 3.3. The pedestrian network represents the walkable street segments in the case study area. Street-level images are collected along this network and linked to the nearest street segments. Multiple images are collected for the same segment so that an average image utility can be calculated. For each segment, the available images are used as visual observations of the walking environment, and the segment is represented by the average predicted image utility. Segments without image data are assigned the median image utility of the corresponding trip-purpose model to prevent disconnected segments.

Route Comparisons

Using a set of origins and destinations on the pedestrian network, the shortest route is first calculated as a reference route. This reference route represents the route that would be selected when only travel time is considered. The route with the highest total utility is then calculated by combining image-based utility with travel time utility. This is done separately for each trip-purpose model, so that route preferences can be compared for walking to public transport, walking to work or school, walking home, and walking in free time.

4

Data Collection and Preparation

This chapter describes how the survey and image data was collected and prepared for analysis. Section 4.1 describes how the survey was implemented and how respondents were recruited. Section 4.2 explains the data cleaning procedure, including the criteria used to exclude low-quality or invalid responses. Section 4.3 describes how the street-level images were prepared for the estimation of the different model specifications. Finally, Section 4.4 explains how the final dataset was divided into training and test data for model estimation and evaluation.

4.1. Survey Data Collection

The SP experiment was implemented as an online survey with the final design shown in Appendix E. The survey was hosted on a TU Delft server, and the data was collected during April 2026. The survey followed the TU Delft Human Research Ethics Committee procedures. Before starting the survey, respondents were informed about the purpose of the study and participated voluntarily. The survey was fully anonymous. No personally identifiable information was collected, only the responses to the survey questions. As a result, individual responses cannot be linked back to respondents, and it is not possible to verify whether a respondent participated more than once.

The target group consisted of respondents who live in the Netherlands, or have lived in the Netherlands for most of their life, and are 18 years or older. Due to practical and budgetary constraints, it was not possible to recruit respondents through a paid survey panel. Instead, the survey was distributed through the researcher's personal and professional network, including LinkedIn, TU Delft students and staff, and contacts at TNO. This means that the respondent pool is based on the researcher's network and is likely not fully representative of the Dutch population. Respondents were also not financially compensated. This should be taken into account when interpreting the results. In total, 310 responses were collected.

4.2. Survey Data Cleaning

Several steps were taken to improve the quality of the collected data. First, incomplete responses were removed. A response was considered complete only if all questions in the survey were answered. During the deployment phase of the survey, a technical issue on some mobile devices caused respondents to only be able to select extreme values for the perception questions. These responses were manually reviewed and removed. Although the stated choice questions in these responses may have been answered seriously and could contain useful information, some respondents indicated that they completed the survey again on a different device. Keeping the original response could therefore lead to the same respondent being included multiple times. Respondents who indicated that they do not currently live in the Netherlands and have not lived in the Netherlands for most of their life were also excluded. Since the study focuses on pedestrian route choice preferences in the Dutch context, these responses were considered outside the target population.

Finally, response times were used as a simple quality check. Extremely fast responses may indicate that respondents did not properly read or consider the questions. A threshold of 3 seconds per survey step or choice task was used as an indicator of very fast responses. Respondents were excluded if they completed four or more steps in under 3 seconds, because this indicates a repeated pattern of very fast responses. This is especially relevant for the choice tasks, as they are presented in groups of three. When a respondent answers multiple tasks too quickly, it is likely that they did not fully process the differences between the route alternatives and trip-purpose contexts.

After applying all filtering steps, 193 valid responses remained for analysis. Each respondent completed 12 route choice tasks, consisting of three tasks for each of the four trip-purpose contexts. This resulted in 2,316 observed choices in total, or 579 observations per trip purpose. The final cleaned sample provides enough observations to estimate the proposed model specifications and to compare patterns in stated route choice preferences across trip-purpose contexts. However, the sample should be interpreted as a convenience sample. The results are therefore used to analyse stated trade-offs within the collected respondent group, rather than to make statistically representative claims about all pedestrians in the Netherlands.

4.3. Image Data Preparation

In addition to the survey response data, the image data also needed to be prepared for model estimation. Only the images that appear in the valid survey responses are used for estimating the models. Images that were collected during the street-level image collection process but were not shown in any valid choice task are excluded from the estimation dataset.

For the pixel-share MNL model, the pixel-share variables extracted during the image analysis and filtering process are used. As described in Section 3.3, these variables are obtained using the Mask2Former segmentation model (Cheng et al., 2022). Before estimation, the pixel-share variables are multiplied by 100, so that they are expressed as percentages rather than proportions. This transformation does not change the information contained in the variables, but only changes the unit in which they enter the utility function. The estimated coefficient then shows the utility change associated with a one percentage point increase in, for example, vegetation share or vehicle share.

For the CV-DCM, the image preparation follows van Cranenburgh and Garrido-Valenzuela (2025). The images are resized and normalised according to the input requirements of the pre-trained DeiT model. This is needed because the model requires a consistent image size and pixel value range. During training, some images are also transformed horizontally. This is a common technique in computer vision to increase variation in the training data and reduce the risk that the model learns image-specific patterns instead of more general visual patterns.

4.4. Train-Test Split

As described in Section 3.5.5, an image-level train-test split is used for model estimation and evaluation. In the final dataset, each street-level image is assigned either to the training set or to the test set, so that the same image does not appear in both sets. A respondent-level split was also explored, but this led to substantial image overlap between the training and test set. A combined respondent- and image-level split was explored using a network-based approach, but this resulted in one connected graph. This means that the data could not be separated in this way. Therefore, the final dataset is split at the image level. The same train-test split is used for all model specifications. This resulted in 1,853 choice observations in the training set and 463 choice observations in the test set. There is no image overlap between the training and test set. Because the split is made at the image level, the same respondent can appear in both sets. The test set therefore evaluates how well the models generalise to unseen images, rather than to unseen respondents.

5

Results

This chapter presents the survey results and the Discrete Choice Models (DCMs) estimated in this thesis. It addresses the final two sub-questions: *How do pedestrian route choice preferences and model performance differ across trip-purpose contexts?* and *How can a visually-enriched route choice model be applied to analyse trip-purpose-specific route preferences across a pedestrian network?*

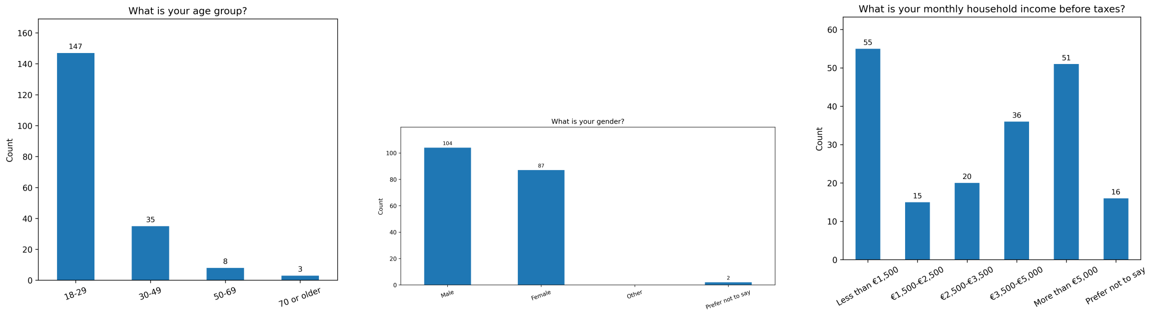
Section 5.1 presents the descriptive survey results and provides information about the collected sample. Section 5.2 then discusses the estimated models, including parameter estimates, the relative importance of attributes, and predictive performance across trip-purpose contexts. The Computer Vision-Enriched Discrete Choice Model (CV-DCM) is examined further through qualitative validation in Section 5.3. Finally, Section 5.4 applies the image-based model to a pedestrian network in Amsterdam-Zuid.

5.1. Descriptive Results

This section discusses the descriptive results of the survey. The goal is to give an overview of the socio-demographic characteristics of the respondents, the general patterns in the choice task data, and the results of the image perception questions.

5.1.1. Socio-Demographics and Walking Habits

The survey was distributed among study groups at TU Delft and personal networks. This means that many respondents were students or young adults. This convenience sample is reflected in Figure 5.1, which shows three main demographic characteristics of the respondent group. Most respondents are between 18 and 29 years old, while older age groups are much less represented. The income distribution shows a similar pattern, with the category below 1500 euros occurring most often.



(a) Age group

(b) Gender

(c) Household income

Figure 5.1: Main demographic characteristics of the respondent group.

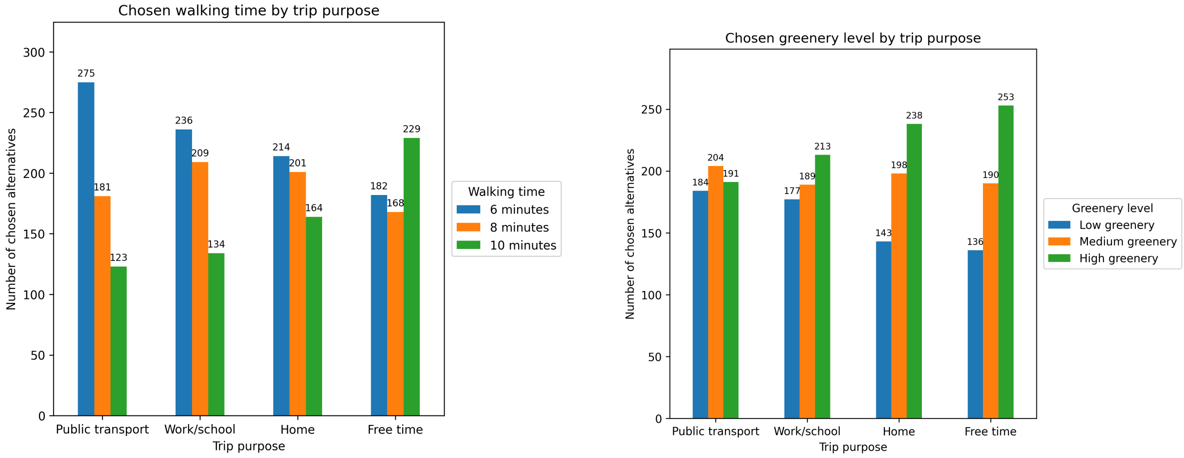
The respondent group can be considered active in walking and generally positive towards walking. Many respondents indicated that they walk almost daily, often in their free time. Figure 5.2 shows the general walking habits of the sample. These results suggest that the respondent group is familiar with walking and could therefore make meaningful choices in the choice tasks based on their own experience. At the same time, the sample is not representative of the Dutch population as a whole. This should be kept in mind when interpreting the results. For this reason, the further analysis focuses on overall patterns in route choice preferences rather than on differences between demographic groups. An overview of the responses to all demographic questions is provided in Appendix F.

(a) Walking frequency **(b)** Main walking purpose **(c)** Attitude towards walking

Figure 5.2: General walking habits of the respondent group.

5.1.2. General Choice Task Patterns

The full model-based analysis is presented in Section 5.2, but some initial patterns can already be seen from the answer distributions. Figure 5.3 shows the distribution of chosen travel times and greenery levels per trip purpose. These two attributes are shown because they show the clearest differences between trip purposes. Since 193 respondents each completed three choice tasks per trip purpose, there are 579 chosen alternatives per trip purpose. The plots for pedestrian presence and vehicle presence are shown in Appendix G.1.



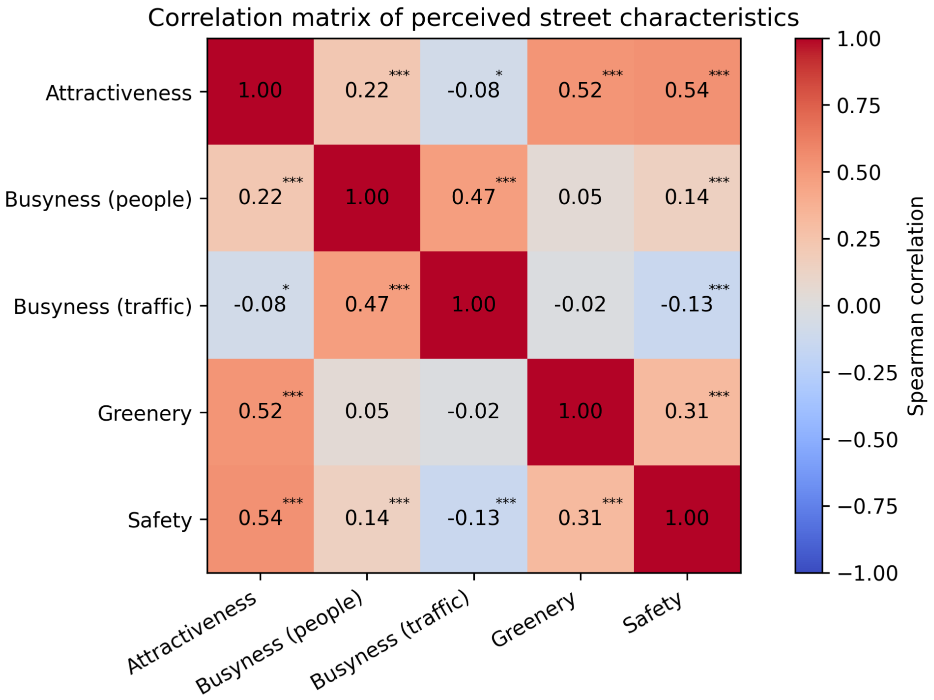
(a) Chosen travel time per trip purpose **(b)** Chosen greenery level per trip purpose

Figure 5.3: Chosen travel time and greenery level per trip purpose.

Figure 5.3a shows that the distribution of chosen travel times differs between trip purposes. For walking to public transport, the shortest travel times are chosen most often. For walking in free time, longer travel times are chosen more often. A similar pattern can be seen for greenery levels. While higher greenery levels are generally chosen more often, it is chosen the most for the walking in free time context. The other image characteristics, pedestrian presence and vehicle presence, also vary across trip purposes, but these patterns are less clear than for travel time and greenery. These results provide an initial indication that trip purpose may influence route choice preferences. However, this still requires further analysis. Each route alternative consists of multiple attributes at the same time, and the image contains more information than the characteristics shown here.

5.1.3. Single-Image Perceptions

The single-image perception answers were used to examine how different perceived characteristics within a single image are related to each other. Figure 5.4 shows the relationships between the five perception scores using Spearman correlations. This correlation measure was used because the perception scores are ordinal. In this figure, the statistical significance of each correlation coefficient is indicated with stars.



	Attractiveness	Busyness (people)	Busyness (traffic)	Greenery	Safety
Attractiveness	1.00	0.22***	-0.08*	0.52***	0.54***
Busyness (people)	0.22***	1.00	0.47***	0.05	0.14***
Busyness (traffic)	-0.08*	0.47***	1.00	-0.02	-0.13***
Greenery	0.52***	0.05	-0.02	1.00	0.31***
Safety	0.54***	0.14***	-0.13***	0.31***	1.00

Figure 5.4: Spearman correlation matrix of perceived street characteristics.
Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The Spearman correlations show the strongest relationships between attractiveness, safety, and greenery. Attractiveness has the highest correlations with safety and greenery. This indicates that streets perceived as attractive are often also perceived as safer and greener. The correlation between greenery and safety is weaker, but still significant. This suggests that these perceptions capture different aspects of the street environment, while both are related to overall attractiveness.

Busyness with people and busyness with traffic are also positively correlated. This means that images perceived as busy with people are often also perceived as busy with traffic. However, these two perceptions relate differently to attractiveness and safety. Busyness with people shows a positive correlation with attractiveness and safety, while busyness with traffic shows a negative correlation with safety. This is in line with the expectation that cars and other motorised traffic can be associated with lower perceived safety.

Overall, the descriptive results suggest that some perception variables are related, especially attractiveness, safety, and greenery. At the same time, none of the correlations are extremely high. The highest correlation is between attractiveness and safety, but it can still be considered moderate. This means that the variables do not appear to be direct proxies for each other. For example, the amount of greenery alone does not directly explain the attractiveness of an image. The perception variables are connected, but they still capture different aspects of how respondents perceived the street images.

5.2. Model Estimation Results

This section addresses the fourth sub-question: *How do pedestrian route choice preferences and model performance change when visual attributes are included, and how do these effects differ across trip-purpose contexts?*

To answer this research question, three model specifications are estimated in this thesis. Model 1 is the basic MNL model and only includes travel time. This model does not use any information from the images and is used as a benchmark model. Model 2 is the Pixel-share MNL model. It includes travel time and visual variables derived from image segmentation, namely sidewalk presence, pedestrian area share, vehicle share, greenery share, and people share. Model 3 is the CV-DCM. This model includes travel time as a numerical attribute and uses the full street-level image as input.

Each model specification is estimated for five model scopes. First, a combined model is estimated using all observations together, without distinguishing between trip purpose contexts. Then, four trip-purpose-specific models are estimated separately for walking to public transport, walking to work or school, walking home, and walking in free time.

$$\textbf{Model 1: } U_{in} = \underbrace{\beta_{\text{time}} \cdot tt_{in}}_{\text{Utility from travel time}} + \epsilon_{in} \tag{5.1}$$

$$\textbf{Model 2: } U_{in} = \underbrace{\beta_{\text{time}} \cdot tt_{in}}_{\text{Utility from travel time}} + \underbrace{\sum_a \beta_a \cdot ps_{ian}}_{\text{Utility from predefined pixel-share attributes}} + \epsilon_{in} \tag{5.2}$$

$$\textbf{Model 3: } U_{in} = \underbrace{\beta_{\text{time}} \cdot tt_{in}}_{\text{Utility from travel time}} + \underbrace{\sum_k \beta_k \cdot z_{ikn}}_{\text{Utility from learned image features}} + \epsilon_{in} \tag{5.3}$$

Here, tt_{in} is the travel time of alternative i for individual n. In Model 2, ps_{ian} represents the pixel share of predefined visual attribute a in the image shown for alternative i to individual n. In Model 3, z_{ikn} represents learned image feature k for alternative i and individual n.

5.2.1. Selected Model Settings

As part of the model optimisation, two alternative specifications were tested for the pixel-share MNL model, as described in Section 5.2.2. First, reduced specifications were tested by including and excluding different pixel-share variables. The changes in overall model performance and statistical significance were small. However, these models did result in lower BIC values, because they included fewer parameters while keeping similar performance. Second, a specification with an alternative-specific constant was estimated. For the trip purpose walking home, this constant was only weakly statistically significant. This suggests that there may be a small general preference for one of the two alternatives in this context, independent of the included route attributes. However, adding the constant did not lead to important changes in the size or direction of the other parameter estimates. This means that the main interpretation of the model remains the same. Therefore, the specification without an alternative-specific constant is preferred, because it is simpler and easier to compare across trip purposes. The results of these alternative pixel-share MNL specifications are reported in Appendix G.2.

For the CV-DCM, hyperparameter tuning first focused on finding suitable settings for the combined model. A broad search was conducted comparing Stochastic Gradient Descent (SGD) and Adam. In this search, SGD performed best. The search was then refined around the best-performing range, which resulted in the selected combined hyperparameter configuration shown in Table 5.1. As described in Section 3.5.5, further optimisation was then performed for the trip-purpose-specific models based on this combined configuration. The hyperparameter tuning results show that different trip-purpose models have different optimal configurations. This means that there is no single optimal setting across all contexts. To keep the models comparable, the configuration that performed best on average across the trip-purpose models is selected. This configuration is also shown in Table 5.1. The estimation results using the best trip-purpose-specific configurations, together with the corresponding hyperparameter settings, are reported in Appendix G.3.

Model	Optimiser	Learning Rate	Batch Size	Weight Decay (γ)
Combined CV-DCM	SGD	$2\mathrm{e}{-5}$	10	0.1
Trip-Purpose Specific CV-DCM	SGD	$2\mathrm{e}{-5}$	8	0.2

Table 5.1: Selected hyperparameter configurations for the CV-DCM

5.2.2. Parameter Estimates

The estimated parameters show how the included variables contribute to route utility in the different model specifications. Table 5.2 presents the parameter estimates and performance values for the three model types and the trip-purpose-specific models. Statistical significance is indicated with significance stars.

When interpreting the estimates, it is important to note that the trip-purpose-specific models are estimated separately. The signs of the coefficients can be compared across models, because they show whether a variable is associated with higher or lower utility within that model. However, the coefficient sizes should not be treated as directly comparable across the different models. This is because each model is estimated on a different subset of the data and may have a different utility scale. For this reason, differences in the strength of the effects are further examined using relative importance metrics.

	Model 1: Basic MNL[I]					**Model 2: Pixel-share MNL[I]**					**Model 3: CV-DCM**				
Walking Purpose	**Combined**	**Public Transport**	**Work or School**	**Home**	**Free Time**	**Combined**	**Public Transport**	**Work or School**	**Home**	**Free Time**	**Combined[II]**	**Public Transport[III]**	**Work or School[III]**	**Home[III]**	**Free Time[III]**
Model information															
Model	RUM-MNL	RUM-MNL	RUM-MNL	RUM-MNL	RUM-MNL	RUM-MNL	RUM-MNL	RUM-MNL	RUM-MNL	RUM-MNL	CV-DCM	CV-DCM	CV-DCM	CV-DCM	CV-DCM
No. param	1	1	1	1	1	6	6	6	6	6	86m	86m	86m	86m	86m
Train set															
N	1853	463	469	464	457	1853	463	469	464	457	1853	463	469	464	457
LL	-1230.69	-269.91	-291.45	-308.60	-314.08	-1169.88	-259.55	-273.65	-290.55	-294.09	-921.45	-150.55	-202.35	-192.91	-232.64
ρ^2	0.042	0.159	0.103	0.040	0.008	0.089	0.191	0.158	0.097	0.072	0.283	0.533	0.376	0.401	0.262
CE	0.664	0.583	0.621	0.665	0.687	0.631	0.561	0.583	0.626	0.644	0.497	0.324	0.432	0.415	0.511
HR	0.604	0.700	0.652	0.591	0.536	0.643	0.706	0.695	0.651	0.630	0.755	0.912	0.844	0.847	0.758
BIC	2468.90	545.97	589.05	623.34	634.29	2384.90	555.94	584.20	617.94	624.93	–	–	–	–	–
Test set															
N	463	116	110	115	122	463	116	110	115	122	463	116	110	115	122
LL	-298.67	-70.58	-64.33	-73.22	-86.34	-276.97	-66.31	-65.01	-64.22	-77.14	-258.27	-62.91	-69.83	-60.96	-77.91
ρ^2	0.069	0.122	0.156	0.081	-0.021	0.137	0.175	0.147	0.194	0.088	0.195	0.204	0.092	0.228	0.093
CE	0.645	0.608	0.585	0.637	0.708	0.598	0.572	0.591	0.558	0.632	0.558	0.552	0.629	0.535	0.628
HR	0.631	0.672	0.682	0.661	0.443	0.680	0.698	0.727	0.730	0.631	0.719	0.746	0.640	0.737	0.710
BIC	603.48	145.92	133.36	151.18	177.47	590.76	161.13	158.22	156.92	183.10	–	–	–	–	–
Parameter estimates															
β_{time}	-0.209***	-0.445***	-0.349***	-0.206***	0.091**	-0.286***	-0.512***	-0.441***	-0.303***	0.024	-0.307	-0.431	-0.320	-0.271	-0.120
β_{side}	–	–	–	–	–	0.001	-0.002	-0.011	0.010	0.000	–	–	–	–	–
β_{area}	–	–	–	–	–	0.011***	0.013*	0.014*	0.013	0.009	–	–	–	–	–
β_{vehi}	–	–	–	–	–	-0.001	0.017	-0.017	-0.007	0.008	–	–	–	–	–
β_{gree}	–	–	–	–	–	0.026***	0.021***	0.026***	0.029***	0.031***	–	–	–	–	–
β_{peop}	–	–	–	–	–	0.037	0.028	0.058	0.046	-0.005	–	–	–	–	–

[I] For Models 1 and 2, using significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
[II] Combined CV-DCM hyperparameters: {optimiser: SGD, learning rate: $2\mathrm{e}{-5}$, batch size: 10, L2 = 0.1}
[III] Trip-purpose-specific CV-DCM hyperparameters: {optimiser: SGD, learning rate = $2\mathrm{e}{-5}$, batch size = 8, L2 = 0.2}

Table 5.2: Performance and parameter comparison across model types and trip purposes.

Model 1: Basic MNL

The basic MNL model only includes travel time as an explanatory variable. In the combined model, the travel time coefficient is negative and statistically significant. This means that, when all trip-purpose data is combined, longer travel times are associated with lower utility. This is in line with the expectation that travel time is generally experienced as a cost in route choice.

When the model is estimated separately by trip purpose, the expected negative sign is found for walking to public transport, walking to work or school, and walking home. The coefficients are negative and statistically significant, which means that longer travel times are associated with lower utility in these contexts. The model for walking in free time shows a different pattern. In this model, the travel time coefficient is positive and statistically significant. This means that, within this model and dataset, longer travel times are associated with higher utility. This may reflect that walking in free time is less time-constrained, so respondents may be more willing to accept or prefer longer routes. However, the experiment included both travel time and street-level images, while the basic MNL model only includes travel time. Therefore, the positive coefficient may also be related to visual or contextual differences between the route alternatives that are not included in this model.

Model 2: Pixel-share MNL

The pixel-share MNL model includes both travel time and pixel-share variables derived from the images. These variables describe the share of pixels in the image representing sidewalks, pedestrian areas, vehicles, greenery, and people. In the combined model, the travel time coefficient remains negative and statistically significant. This means that longer travel time is still associated with lower utility after adding the image variables. In the trip-purpose-specific models, travel time is negative and statistically significant for walking to public transport, walking to work or school, and walking home. For walking in free time, the coefficient is positive but no longer statistically significant. This differs from the basic MNL model, where the travel time coefficient was positive and significant for walking in free time.

Of all pixel-share parameters, greenery share shows the clearest and most consistent effect. The coefficient is positive and statistically significant in the combined model and in all trip-purpose-specific models. This means that the streets with more visible greenery are associated with higher utility across all walking contexts. This is in line with the expectation that pedestrians generally prefer greener walking environments. Pedestrian area share is also positive in all models. It is statistically significant in the combined model, for walking to public transport, and for walking to work or school, but not for walking home and walking in free time. This suggests that pedestrian-oriented infrastructure may also be associated with higher utility, although this effect is less consistent than the effect of greenery.

Some of the remaining signs are less in line with the expected behavioural direction. For example, sidewalk presence has different signs across trip purposes. The expectation would be that it has a similar direction as pedestrian area share, because both variables relate to pedestrian space. However, the sidewalk coefficients are not statistically significant in any of the models. Their signs should therefore not be interpreted as clear behavioural effects. This does not mean that sidewalk presence is irrelevant, but rather that this model does not provide enough statistical evidence to estimate a clear effect.

To assess the relative importance of the included variables within each model, the potential utility change of each attribute is calculated, as described in Section 3.6.1. The percentages in Table 5.3 show how much each attribute contributes to the maximum possible utility difference within each model.

Attribute	**Combined**	**Public transport**	**Work or school**	**Going home**	**Free time**
Travel time	23%	39%	26%	23%	3%
Sidewalk presence	$< 1\%$	2%	6%	7%	$< 1\%$
Pedestrian area share	12%	12%	11%	12%	12%
Vehicles	$< 1\%$	10%	8%	4%	7%
Vegetation share	45%	28%	28%	41%	74%
People share	19%	9%	21%	13%	4%

Table 5.3: Contribution to the maximum possible utility difference in the Pixel-share MNL models.

For walking in free time, greenery share has the largest contribution, with 75%, while travel time only accounts for 3%. This suggests that, in this context, utility differences are mainly related to changes in greenery, while changes in travel time play a much smaller role. For walking to public transport, travel time has the largest contribution, with 39%. This fits with the expectation that this context is more time-sensitive. Pedestrian area share is more stable across the models, with a contribution of around 12%. Although the coefficient is not statistically significant in all models, this may suggest that the share of pedestrian area in an image is valued more consistently across trip purposes.

Model 3: CV-DCM

Only the travel time coefficient is reported for the CV-DCM in Table 5.2. This is because the model contains a large number of parameters related to the image representation, which are not directly interpretable. The analysis therefore focuses on the travel time coefficient and on the relative contribution of travel time and image utility to the maximum possible utility differences.

Overall, the travel time coefficient in the CV-DCM is negative across all trip purposes. This means that longer travel times are associated with lower utility in each context, including walking in free time. This differs from the two MNL models, where the travel time coefficient was positive for walking in free time. This suggests that the effect of travel time may depend on how visual information is included in the model. In the MNL models, the positive coefficient showed that respondents are more willing to accept longer walks in their free time. In the CV-DCM, part of this effect may instead be captured by the image utility, because the model uses the full street-level image as input.

The relative importance values in Table 5.4 show that image utility has the largest contribution in all CV-DCM models. This means that most of the maximum possible utility difference comes from visual information learned from the images. This is especially clear for walking in free time, where image utility accounts for 90% and travel time accounts for 10%. This is in line with the results of the basic and pixel-share MNL models, where travel time showed less significance for this trip purpose. It suggests that the information from the images might be able to explain the route choice preferences better than than travel time.

For walking to public transport and walking to work or school, travel time has a larger contribution than for walking in free time, with 21% and 22%, respectively. This fits with the expectation that these contexts are more time-sensitive. However, image utility still accounts for most of the relative importance in these models. This suggests that the images contain relevant information for explaining the choices in the dataset across all trip purposes.

Attribute	**Combined**	**Public transport**	**Work or school**	**Going home**	**Free time**
Travel time	21%	21%	22%	17%	10%
Image utility	79%	79%	78%	83%	90%

Table 5.4: Contribution to the maximum possible utility difference in the CV-DCM models.

5.2.3. Predictive Performance

The predictive performance of the models is compared using log-likelihood, cross-entropy, hit-rate, ρ^2, and BIC where available. These metrics are described in Section 3.6.2 and reported in Table 5.2. All models are estimated by maximising the log-likelihood or by minimising the cross-entropy. Overall, the results show that adding visual information improves model performance compared with the basic MNL model in the combined model and in most trip-purpose-specific models. For the combined test set, the cross-entropy decreases from 0.645 for the basic MNL, to 0.598 for the pixel-share MNL, and to 0.558 for the CV-DCM. The hit-rate also increases from 0.631 to 0.680 and 0.719. This shows that both the predefined pixel-share variables and the learned image features add information that helps explain the observed choices. Because the results differ across trip purposes, the rest of this section focuses mainly on the trip-purpose-specific models.

For the pixel-share MNL, the improvement compared with the basic MNL is visible in most trip-purpose contexts. This shows that even a relatively simple representation of the image can improve predictive performance. The largest improvements are found for walking home and walking in free time. For walking home, the test cross-entropy decreases from 0.637 to 0.558, while the hit-rate increases from 0.661 to 0.730. For walking in free time, the hit-rate increases from 0.443 to 0.631. This is in line with the expectation that visual street characteristics are more relevant for less time-constrained trips. However, the BIC values of the trip-purpose-specific pixel-share models are higher than those of the basic MNL. This means that, for these separate contexts, the improvement in model fit is not always large enough to justify the additional parameters. The optimisation results in Appendix G.2 show that removing less informative pixel-share variables can improve the balance between model fit and model complexity.

The CV-DCM shows the strongest improvement in the training data. For example, for walking home, the training ρ^2 increases from 0.040 for the basic MNL to 0.401 for the CV-DCM. The training hit-rates are also high, with values above 0.75 for all contexts and 0.912 for walking to public transport. This large difference is expected to some extent, because the CV-DCM has a much more flexible model structure than the two MNL-based models. It is not limited to predefined image variables and can learn more complex visual patterns from the images. At the same time, the difference between the training and test results shows that the model fits the training data much more strongly than the test data. This suggests a risk of overfitting. However, this does not mean that the model is unusable. The test results still outperform the other models in most contexts, which indicates that the model also learns visual information that generalises to unseen images.

When comparing the test results across trip purposes, the largest improvements are found for walking home and walking in free time. This is most visible in the hit-rates. For walking in free time, the CV-DCM improves the hit-rate from 0.443 to 0.710. For walking home, it increases from 0.661 to 0.737. These results fit the expectation that visual street characteristics become more relevant when the trip is less goal-oriented. However, this is not the full explanation, because the public transport context also performs better when using the CV-DCM. For walking to public transport, the CV-DCM has the lowest test cross-entropy of 0.552 and the highest hit-rate of 0.746. This suggests that visual information may also matter in more goal-oriented contexts, although the type of visual information that matters may differ by trip purpose.

One clear exception is the walking to work or school context. In this context, the basic MNL performs best on most test metrics in Table 5.2. It has a test cross-entropy of 0.585 and a ρ^2 of 0.156, while the CV-DCM has a cross-entropy of 0.629 and a ρ^2 of 0.092. The pixel-share MNL performs close to the basic MNL, and even has a higher hit-rate of 0.727 compared with 0.682. This suggests that, under the shared hyperparameter configuration used in Table 5.2, adding visual information does not clearly improve predictive performance for the context of walking to work or school. This is partly in line with the expectation that these trips are more time-driven. However, the result should be interpreted carefully. The optimised configurations in Appendix G.3 show that both the pixel-share MNL and the CV-DCM can outperform the basic MNL for this context when the model settings are selected specifically for this trip purpose. This indicates that the weaker performance in Table 5.2 is not necessarily evidence that visual information is irrelevant for work or school trips. Instead, it shows that the results are sensitive to the selected model specification and training settings.

Overall, the predictive performance results show that visual information can improve model fit and prediction, but that this improvement differs between trip purposes and model settings. Since the CV-DCM generally shows the strongest predictive performance, the next step is to inspect whether its predicted image utilities are also meaningful from a visual and behavioural perspective.

5.3. Qualitative Model Validation

To further interpret the CV-DCM, the predicted image utilities are inspected qualitatively. The purpose of this validation is to assess whether images with high predicted utility also look like attractive walking environments, and whether images with low predicted utility contain characteristics that are expected to reduce walking preference. The validation also helps identify cases where the model behaves less as expected. This is important because the CV-DCM can learn complex visual patterns, but these patterns are not always easy to interpret.

The qualitative validation is structured in four steps. First, the predicted image utilities are compared with the image-level perception scores. Second, the highest- and lowest-utility images are inspected for each trip purpose. Third, images from the same street are compared visually and based on their differences in utility. Finally, images edited using generative AI are used to test how the model responds to changes in greenery, traffic, trash, and benches.

5.3.1. Comparison with Perception Scores

The predicted image utilities are compared with the perception scores to determine whether images with higher utility are also perceived more positively. This comparison is only possible for images that were both used in the choice tasks and rated in the perception questions. Because the number of rated images is limited compared to the full image set, not all choice-task images have a perception score. Each rated image was also evaluated by only one respondent. The results should therefore be interpreted as an indication of the relationship between image utility and perceived street quality, rather than as a full validation.

Perception	Combined		Public transport		Work/school		Home		Free time	
	Top	Bottom	Top	Bottom	Top	Bottom	Top	Bottom	Top	Bottom
Attractiveness	3.65	1.70	3.55	1.95	3.50	1.90	3.45	1.65	3.85	2.15
Greenery	3.70	1.30	3.40	1.70	3.00	1.30	3.05	1.60	4.00	1.45
Safety	3.35	2.85	3.75	2.65	3.50	2.90	3.30	2.70	3.25	2.65
Busyness people	2.25	1.90	2.00	2.10	1.90	1.85	1.70	1.65	1.70	1.80
Busyness traffic	2.10	2.30	1.90	2.20	2.00	2.15	1.65	2.70	1.55	2.30

Table 5.5: Average perception scores of the top and bottom 20 image utilities by model

Table 5.5 shows that images with higher predicted utility generally receive higher scores for attractiveness and greenery than images with lower predicted utility. This pattern is visible in all model specifications. For attractiveness, the largest difference is found in the combined model, where the top images score 3.65 and the bottom images score 1.70. For greenery, the difference is especially clear in the combined and free-time models. In the free-time model, the top images score 4.00, compared with 1.45 for the bottom images. For safety, the differences are smaller, but images with higher predicted utility are still rated as safer in all models. This suggests that perceived safety may also be related to image utility, although this relationship is less clear than for attractiveness and greenery.

The busyness scores show a less consistent pattern. For traffic busyness, images with lower predicted utility generally receive higher scores. This is expected, because streets with more traffic are usually less attractive for walking. The clearest difference is found in the free-time model, where the top images score 1.55 and the bottom images score 2.30. For busyness with people, the differences are small and do not show a clear direction. This suggests that the predicted utility is not strongly related to the number of people visible in the image.

Overall, these results suggest that images with higher predicted utility are mainly associated with higher attractiveness and more greenery. They are also associated with higher safety and lower traffic busyness, but to a lesser extent. However, predicted image utility and perception scores are not the same. The perception scores describe how individual images were rated on separate aspects, while image utility is estimated from route choices.

5.3.2. High- and Low-Utility Images

The next validation step is to inspect the images with the highest and lowest predicted utilities. These utilities show which walking environments are most and least preferred by the model within a specific trip-purpose context. Because the models were estimated using route choices that included both travel time and street-level images, the predicted image utilities reflect the visual part of the route preference within that context. For each trip-purpose model, the model is applied to the same set of survey images. This makes it possible to compare whether the most and least preferred walking environments differ between trip purposes. Figure 5.5 shows the three images with the highest and lowest utility for each trip purpose.

Figure 5.5: Highest and lowest predicted image utilities by trip purpose.

The images with the highest utility show several patterns that are in line with behavioural expectations. Most of these images contain visible greenery, trees, clear walking space, or a relatively calm street environment. This is especially visible for walking home and walking in free time. This supports the earlier model results, where greenery was found to have a positive effect on route utility. The images with the lowest utility are visually different from the highest ones. They generally contain less greenery and often show more infrastructure. Several of these images show large roads, bridges, walls, or areas that appear less comfortable for walking. This is also in line with expectations, because these environments are less likely to be perceived as attractive pedestrian routes.

At the same time, the results also show that the model does not rank images only based on simple visual rules, such as more greenery is always better or more traffic is always worse. Some rankings are less intuitive from the researcher's perspective. For example, for walking home, one of the highest-ranked images contains a road with a car and does not show a very clear pedestrian path. This seems unexpected, because traffic presence and limited pedestrian space would normally be expected to reduce utility. Similarly, some lower-ranked images do not necessarily show the least attractive walking environments. They show designated pedestrian paths and are even close to water, which are both factors associated with higher perceived walkability. This indicates that the model may have learned other visual patterns that are less straightforward to explain.

The differences between trip purposes are also important, because the images with the highest utilities differ between models. This suggests that the images are evaluated differently depending on the trip context. The highest-ranked images for walking in free time and walking home appear greener. For walking to public transport and walking to work or school, the highest-ranked images also show greenery, but they include clearer road infrastructure or clear walking space. This fits the expectation that visual quality may matter differently depending on the purpose of the walk. More goal-oriented trips may value clear paths and directness more, while free-time and home trips may be more sensitive to an attractive environment.

Disagreement Between Trip Purpose Models

To further analyse the differences between the trip-purpose models, the images with the largest rank differences are shown in Figure 5.6. This shows where the models disagree most strongly about the same image. Some images are ranked very differently depending on the trip purpose, which suggests that the same street environment is not evaluated in the same way across walking contexts.

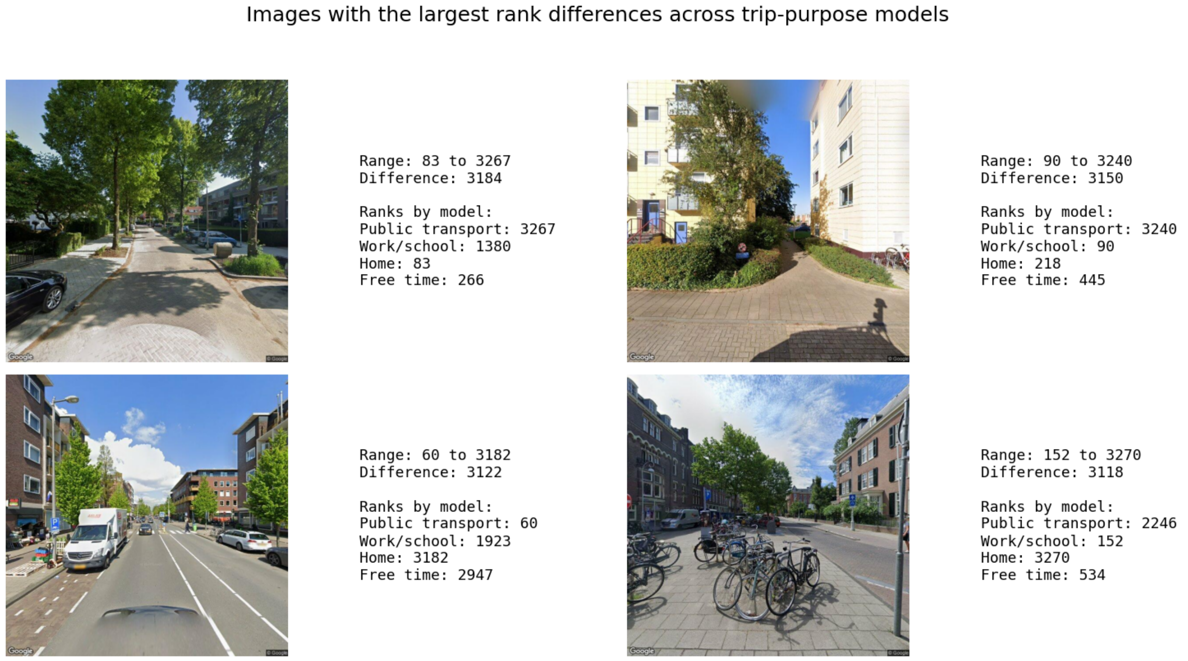


Figure 5.6: Images with the largest rank differences across trip-purpose models.

The greener images are ranked among the highest for walking home or walking in free time, but among the lowest for walking to public transport. The opposite pattern is also visible. The image with wider roads, more urban infrastructure, and more traffic is ranked among the highest for walking to public transport, but much lower for the other trip purposes. These large rank differences support the use of separate trip-purpose models. They suggest that the models may learn trip-purpose-specific relationships, instead of simply assigning high utilities to greener images.

5.3.3. Same-Street Image Comparison

The same-street comparison is used to look at differences in utility within the models. The idea is that images taken only a few metres apart would be expected to receive relatively similar utilities. Figure 5.7 shows the same-street image comparison. The contexts of walking to work and school and free time are not included here, since the other two models showed more interesting results. These examples are shown in Appendix G.4.

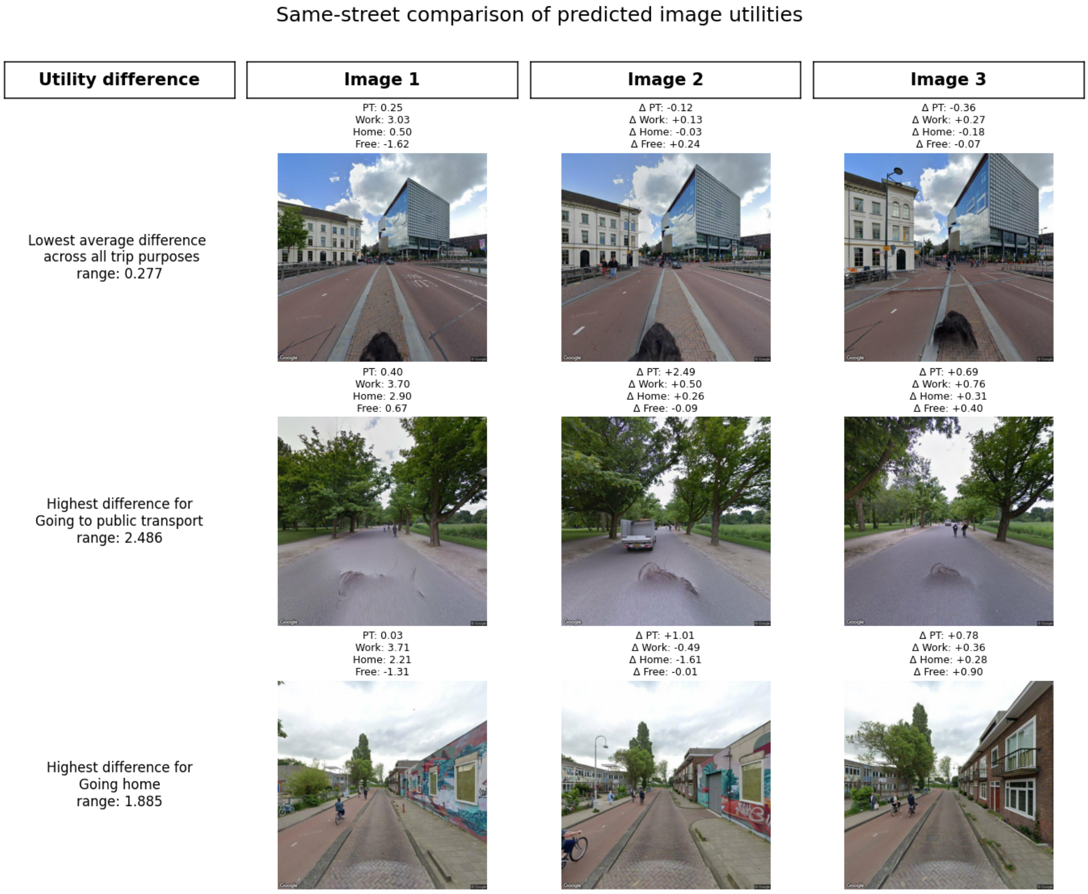


Figure 5.7: Same-street comparison of predicted image utilities.

The first row shows the street with the lowest average utility difference across all trip purposes. The predicted utilities are similar for the three images, which is expected because the images show almost the same street environment. The small differences are likely caused by changes in perspective. For example, the amount of visible buildings and road space changes slightly between the images. This suggests that, in this case, the models assign relatively stable utilities when the visual environment remains mostly unchanged.

The second and third rows show larger differences within the same models. In the public transport model, the image with a visible vehicle receives the highest utility. This is somewhat unexpected, because vehicle presence would usually be associated with lower street attractiveness. One possible explanation is that the public transport model has learned that images with vehicles were more often preferred in the stated choice data for this trip purpose. For the walking home model, the second image is the outlier with a much lower utility. A possible explanation is that the pedestrian path appears to be interrupted. Interestingly, the first and third images have similar utilities, even though they differ visually. One image shows graffiti on the wall, while the other does not.

Overall, the comparison shows that predicted utilities are relatively stable when images are very similar. However, larger differences can occur within the same model, and these differences are not always explained by one clear visual element.

5.3.4. Image Augmentation Results

As a final exploratory validation, the model is tested using versions of the same images edited with generative AI. Four types of edits are made: adding more greenery, adding more traffic, adding visible trash, and adding benches. Some of the edited images are shown in Figure 5.8. The full set of edited images and the exact prompts used to generate them are shown in Appendix G.5.

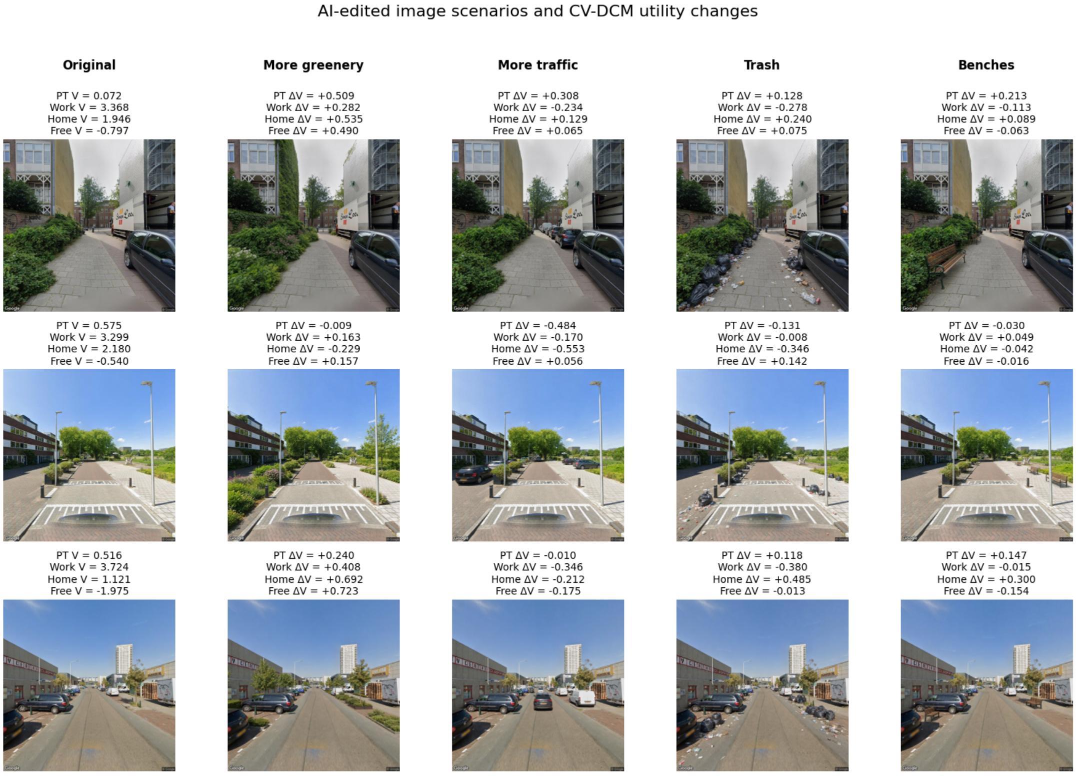


Figure 5.8: AI-edited image scenarios and changes in predicted CV-DCM image utility.

For these three edited images, adding more greenery generally improves utility. Adding more traffic shows a weaker and less consistent pattern. For the first image, utility increases, while for the other images utility decreases. For visible trash and benches, the results are even less consistent. In both cases, utility can either increase or decrease. To summarise these effects more clearly, Figure 5.9 shows the utility changes for each image augmentation per trip purpose, based on eight edited images.

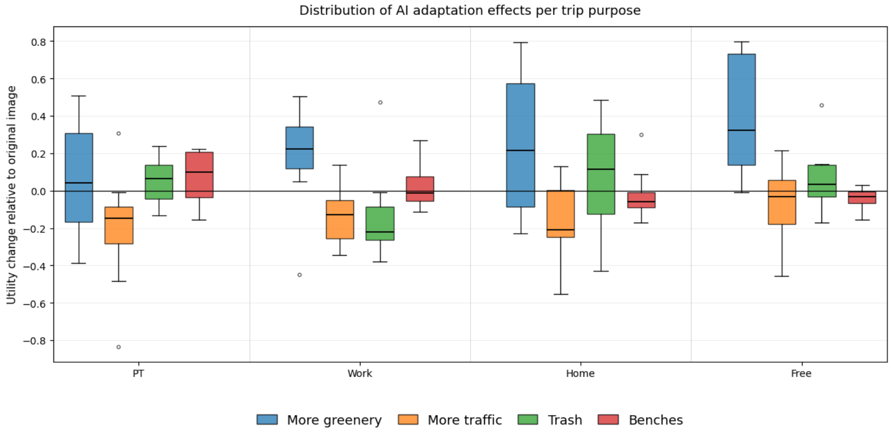


Figure 5.9: Distribution of AI-edited image utility changes per trip purpose for eight edited images.

Overall, the boxplots show that adding more greenery generally increases utility, while adding more traffic generally decreases utility. Adding trash and benches gives less clear results, because the utility changes can be either positive or negative.

For the public transport model, adding more greenery does not always lead to higher utility. Adding more traffic generally leads to lower utility, but there is also an outlier around 0.3. This outlier is in line with the same-street comparison, where the car in the image also seemed to increase utility instead of decreasing it. For the free-time model, adding traffic would be expected to decrease utility, but this effect is not clearly visible in the boxplot.

These inconsistent results do not necessarily mean that the models are wrong. Instead, they show that the augmentations need to be interpreted carefully, because visual characteristics in an image can be connected to each other. Editing one object into an image does not mean that the utility change is directly caused by that object. For example, in the second image in Figure 5.8, adding greenery results in a utility change of -0.229 for the walking home model. However, when looking more closely at the image, the added greenery covers a large part of the sidewalk on the left side. This means that adding one visual element can also remove or hide other aspects of the image, which may be more important in a specific trip context.

5.4. Spatial Application of the CV-DCM

This section presents the results of the spatial application and addresses the final sub-question: *How can a visually enriched route choice model be applied to analyse trip-purpose-specific route preferences across a pedestrian network?* The trip-purpose-specific CV-DCM models are applied to a case study area in Amsterdam-Zuid. The area of interest is shown in Figure 5.10.

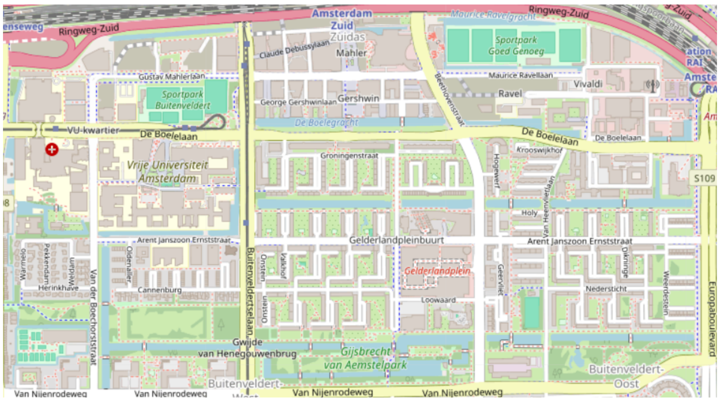


Figure 5.10: Case study area in Amsterdam-Zuid used for the spatial application

For this application, 4,988 street-level images were collected on the pedestrian network. The network contains 3,450 edges, of which 2,126 could be linked to at least one image. Edges without available images were assigned the median image utility within the corresponding trip-purpose model.

5.4.1. Spatial Utility Distribution

Using the collected street-level images, the average image utility is calculated for each street segment and for each trip-purpose model. The resulting spatial distribution is shown in Figure 5.11. In this figure, red indicates the lowest image utilities and green indicates the highest image utilities within each trip-purpose model. The utility range differs between the models, so the colours should mainly be interpreted as the relative utility ranking within each model. Since the same set of street-level images is used for all trip-purpose models, the maps can still be compared to identify similar spatial patterns across trip purposes.

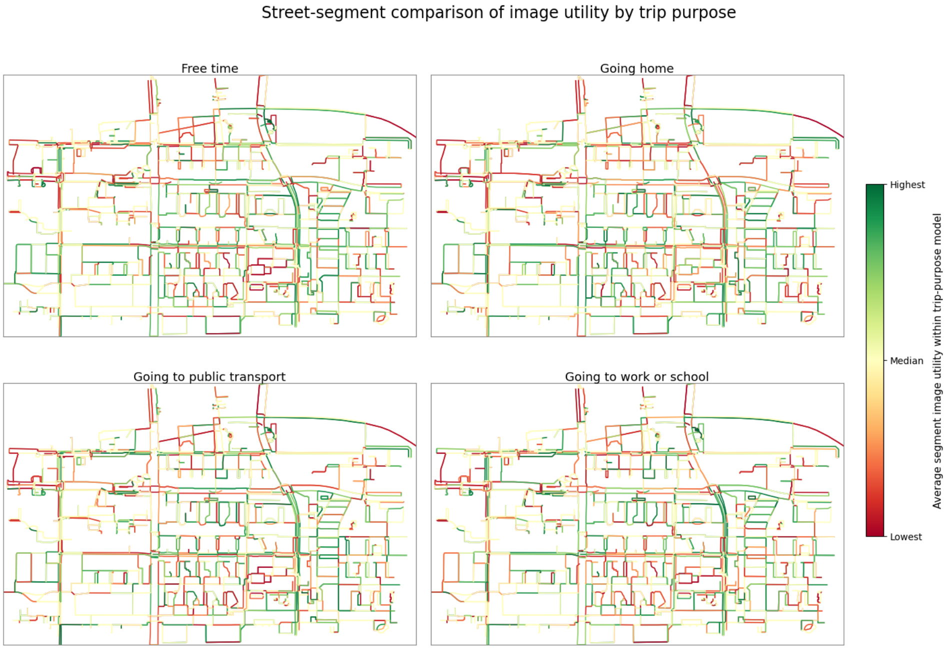


Figure 5.11: Average street-segment image utility in Amsterdam-Zuid by trip-purpose model

Figure 5.11 shows several similarities between the four trip-purpose maps. Across all trip purposes, lower image utilities are mostly found in the upper-left, upper-middle, and lower-middle parts of the map. These areas correspond to locations around Amsterdam UMC, Amsterdam Zuid station, and further south around Gelderlandplein. These locations are relatively busy and more strongly focused on car traffic or have more high-rise buildings. They also contain less visible greenery. This is in line with the earlier model validation results, where images with more infrastructure, cars, and roads generally had lower utilities.

Higher image utilities are more visible in the lower-left part and on the eastern side of the map. The lower-left area includes a neighbourhood garden and the Gijsbrecht van Aemstelpark. These streets are located around greener areas and parks. On the eastern side, around the Rotterdamsepad, the streets are mainly surrounded by apartment buildings, but they also include a relatively high amount of greenery and some water. This suggests that residential streets where buildings are combined with visible green space can also receive high predicted utilities.

The two example locations in Figure 5.12 show this visual difference. The area around Amsterdam UMC mainly shows grey infrastructure and car roads, while the area around Rotterdamsepad shows a greener environment that is not focused on larger motorised traffic. These visual differences align with the previous results, where greener areas generally had higher utilities and infrastructure-dense areas had lower utilities.

(a) Lower utility area around Amsterdam UMC

(b) Higher utility area around Rotterdamsepad

Figure 5.12: Example street environments with lower and higher image utilities in the spatial application

5.4.2. Route Comparisons

Based on the maps described above, four locations in the case study area were selected to create routes between them. The selected routes connect Amsterdam Zuid station to Vrije Universiteit Amsterdam, Vrije Universiteit Amsterdam to the Gijsbrecht van Aemstelpark, and the Gijsbrecht van Aemstelpark to Gelderlandplein. These routes all have a travel time of approximately 8 to 10 minutes when the shortest path is used. They also cross different types of street environments, including station areas, university areas, residential streets, and greener park areas. All route comparisons are shown in Appendix G.6. Figure 5.13 shows the route from Vrije Universiteit Amsterdam to the Gijsbrecht van Aemstelpark.

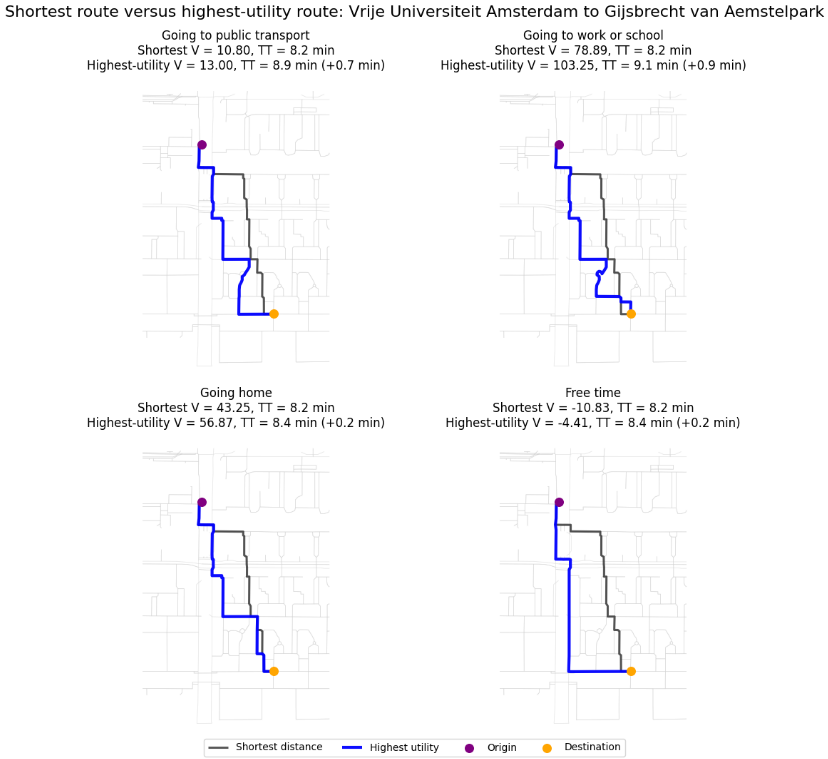


Figure 5.13: Comparison between the shortest route and the highest-utility route from Vrije Universiteit Amsterdam to the Gijsbrecht van Aemstelpark by trip-purpose model.

Figure 5.13 shows that the routes selected by the model differ from the shortest route for all four trip-purpose models. This indicates that route preferences are not only determined by travel time. In this example, the image utilities are large enough to change the preferred route from the shortest route to a visually more attractive route. The preferred routes also differ between trip purposes. The free-time model clearly follows a different route, while the other models follow a similar route at the start but differ closer to the destination. This shows that route preferences can depend on the walking context.

A more unexpected result is that the more goal-oriented models seem to prefer routes with a larger increase in travel time than the other models. This is surprising, because these trip purposes are expected to value shorter routes more strongly. In this case, it seems that the visual characteristics of the preferred routes increase the utility enough to compensate for the extra travel time.

Overall, the route comparison shows that adding visual information to a route choice model can lead to different preferred routes compared with a shortest-route approach. It also shows that these route preferences can differ between trip purposes. However, the results should be interpreted carefully. This is only a small example based on a limited number of images, and the models were trained with survey data from a convenience sample. These routes should not be seen as evidence of actual pedestrian behaviour in this area. Instead, they show how visual street characteristics can influence model-based route preferences.

6

Discussion

This chapter discusses the main findings of this thesis in relation to the research objective, the existing literature, and the methodological choices made in the study. The thesis compared a basic Multinomial Logit (MNL) model, a pixel-share MNL model with predefined visual attributes, and a Computer Vision-Enriched Discrete Choice Model (CV-DCM) with learned image representations across four trip-purpose contexts. By doing so, this research examines how different levels of visual information can be included in models of pedestrian route choice preferences, and how trip-purpose-specific model specifications can lead to different results.

Section 6.1 interprets the main findings and connects them to the literature on pedestrian route choice, walkability, trip purpose, and machine learning in choice modelling. Section 6.2 then discusses the strengths and limitations of the research, focusing on the comparison of model specifications, trip-purpose-specific modelling, validation beyond model fit, sample generalisability, hypothetical bias, and the simplified representation of the walking environment. Section 6.3 discusses the implications of the findings for pedestrian route choice modelling, urban planning and policy, and the responsible use of AI in behavioural modelling and policy. Finally, Section 6.4 provides practical recommendations and directions for future research.

6.1. Interpretation of Main Findings

This section discusses the main findings of the thesis and explains how the results can be interpreted in relation to the field of pedestrian route choice. The interpretation is structured around five key themes: the role of visual street-level information, the effect of greenery as a visual street characteristic, the effect of trip purpose as context, the difference in route preferences on a network, and overall model performance and interpretability. Together, these themes show that visual information can contribute to pedestrian route choice modelling and that route preferences differ across trip-purpose contexts. At the same time, the results also show that some findings are difficult to explain, especially when the image-based models show route preferences that do not fully align with behavioural expectations or literature.

6.1.1. Prediction of Route Choice Preferences with Visual Information

The first important finding is that adding street-level visual information improves the prediction performance of pedestrian route choice preferences for the SP dataset used in this thesis. Both the pixel-share MNL models and the CV-DCM generally outperform the basic MNL model that only includes travel time. This indicates that travel time alone does not sufficiently explain the observed route choices, and that visual information contributes meaningfully to explaining these choices.

This finding is consistent with pedestrian route choice and walkability literature, which shows that pedestrians do not only evaluate routes based on travel time or distance, but also consider environmental qualities such as comfort, perceived safety, greenery, traffic exposure, and general attractiveness (Basu et al., 2022; Tong and Bode, 2022; Sevtsuk et al., 2021; Choi and Kang, 2025). Since travel time alone does not sufficiently explain the observed choices, the results of this thesis support the argument that pedestrian route choice models benefit from including variables that represent the visual qualities of the walking environment.

This finding also relates to the broader discussion on using machine learning in choice modelling. Van Cranenburgh et al. (2022) argue that machine learning can broaden the scope of choice modelling by making it possible to use unstructured data, such as images, text, and video, which are difficult to include directly in traditional choice models. The results of this thesis support this argument, since the CV-DCM uses the full image as input and generally improves prediction performance compared with the basic and pixel-share MNL. This suggests that the model appears to capture visual patterns that are relevant for explaining route choice preferences that could not be fully represented by predefined attributes.

6.1.2. Preference for Greener Walking Environments

Across the different model specifications, greenery stands out as the most robust and consistently positive visual characteristic. In the pixel-share MNL models, the share of visible greenery is positively associated with route utility and statistically significant across all trip-purpose contexts. The qualitative analysis of the CV-DCMs supports this result, as images with higher utility often show greener street environments. The comparison with perception scores also shows that higher-utility images tend to receive higher attractiveness and greenery scores.

This strong preference for greenery is consistent with literature focused on walkability, which shows that greener streets and parks contribute positively to perceived walkability and walking experience (Borst et al., 2008; Choi and Kang, 2025). The results of this thesis are therefore behaviourally plausible, since they align with the idea that pedestrians value greener environments and choose routes not only based on travel time, but also based on perceived environmental quality.

At the same time, this finding differs from the pedestrian route choice model by Sevtsuk et al. (2021), in which greenery was not found to be a statistically significant predictor in their final specification. This difference may be related to how greenery is measured, the model specification, or the type of data used for estimation. In particular, Sevtsuk et al. (2021) use Revealed Preference (RP) data in the form of GPS trajectories, while this thesis uses an Stated Preference (SP) experiment in which respondents evaluate street-level images together with travel time. Jin et al. (2025) show that stated pedestrian route preferences can change when respondents physically walk the routes, because the effort associated with longer distances may be underestimated in hypothetical choices. Actual route choices may also depend on familiarity, daily habits, and cultural context (Duives, 2025; Levine and Norenzayan, 1999). The strong effect of greenery found in this thesis should be interpreted as a preference for greener walking environments, while its influence on actual route choice may be smaller and may also depend on other conditions experienced while walking.

6.1.3. Differences Between Trip Purposes

The results show that pedestrian route choice preferences differ across trip-purpose contexts. By estimating separate models for walking to public transport, walking to work or school, walking home, and walking in free time, it becomes clear that travel time and visual attributes are valued differently across walking contexts. For walking in free time, visual attributes, particularly greenery, explain a large share of the potential utility difference, while the contribution of travel time becomes relatively small. In contrast, for walking to public transport, travel time remains more important, although visual information still contributes to route utility. The public transport and work or school models are very similar, which can be expected because they both represent more goal-oriented walking contexts. The walking home model is more in between the goal-oriented models and the free time model. In this context, travel time still remains important, but the environment appears to play a somewhat larger role.

These findings are consistent with research describing pedestrian route choice as context dependent (Hoogendoorn and Bovy, 2004; Tong and Bode, 2022). Trip purpose does not change the street environment, but it changes the situation in which the route is evaluated. The results for walking to public transport and walking to work or school align with studies showing that more goal-oriented walking is strongly influenced by time and distance, but cannot be explained by these factors alone (Hatamzadeh et al., 2014). The walking home context also differs from the other contexts, which aligns with research arguing that returning home should be considered as a separate walking purpose (Watson et al., 2021). The walking in free time model aligns with literature on recreational walking, where trips are less focused on travel time and where the walking environment plays a more important role (Larranaga and Cybis, 2014).

The differences in predictive performance and in the importance of visual qualities between the trip-purpose models show that pedestrian route choice preferences should not be treated as the same across all walking contexts. The same route can be evaluated differently depending on the purpose of the walk. This contributes to the recent discussion by Zafri and Sevtsuk (2026). They point out that many existing pedestrian models represent pedestrian behaviour in a general way, despite differences between people and walking contexts. The results of this thesis show that there are indeed differences in both the estimated parameters and the routes of the different trip-purpose models. Simplifying pedestrian route choice by assuming one general route choice model would therefore miss these context-specific differences in how pedestrians evaluate and choose routes.

6.1.4. Differences Between Shortest Route and Preferred Routes

The spatial application in Amsterdam-Zuid shows that the preferred routes differ from the shortest routes when visual information is included. When routes are evaluated using both travel time and street-level images, the routes with the highest utility are often longer than the objectively shortest route. This is consistent with pedestrian route choice literature showing that pedestrians may accept a longer route when it offers a more comfortable, safe, or attractive walking environment (Prashker and Bekhor, 2004; Prato, 2009; Sevtsuk et al., 2021; Basu et al., 2023). Within the models estimated in this thesis, the positive utility of a more attractive street can compensate for the negative utility of additional travel time.

The preferred routes also differ between trip-purpose contexts. This means that the differences between the estimated models are large enough to produce different preferred routes for the same origin and destination. At the same time, some of the preferred routes for the more goal-oriented purposes show relatively large increases in travel time, despite travel time being expected to be more important in these contexts (Hatamzadeh et al., 2014). This may also relate to the earlier discussion on differences between stated preferences and actual walking behaviour. Because respondents evaluated hypothetical routes without physically experiencing the additional walking time, the perceived cost of choosing a longer route may have been underestimated (Jin et al., 2025).

6.1.5. Model Interpretability

The results highlight a clear trade-off between predictive performance and interpretability when incorporating visual street-level characteristics into pedestrian route choice models. The basic MNL is easy to interpret because route utility is explained only through travel time, but it has lower predictive performance. The pixel-share MNL obtains better predictive performance by adding predefined visual attributes directly to route utility. However, it still simplifies the street environment into a limited set of visual attributes. The CV-DCM generally performs best in terms of predictive performance across most trip-purpose contexts, indicating that the full street-level images contain relevant visual information for explaining the route choice preferences observed in this study. However, because the image representation is learned by the model, it is less clear which exact visual characteristics increase or decrease the utility.

This reduced interpretability becomes particularly relevant when model outcomes do not fully align with behavioural expectations. For example, in the same-street comparison, the public transport model shows an increase in image utility when a car is visible in the image. However, this comparison only considers the utility of the images and does not take into account the travel times under which the original choices were made. This does not imply that pedestrians prefer seeing cars when walking to public transport, but rather suggests that the CV-DCM has learned associations in the data where images containing a car were chosen more often in this trip-purpose context. Together with the unexpected route rankings found in the spatial application, this shows that the outcomes of the CV-DCM are not always easy to explain behaviourally.

This trade-off in interpretability highlights the other side of the discussion of using machine learning in choice modelling. Van Cranenburgh et al. (2022) argue that behavioural interpretation becomes more difficult when model parameters do not have a direct behavioural meaning. A similar point is made by Zhu and Si (2024), who find that DCMs can achieve similar predictive performance to deep neural network models, while being easier to explain. This suggests that more complexity does not always imply better performance and that the machine learning models should not automatically replace the more interpretable DCMs. The most suitable model depends on the purpose and context of the study, especially when interpretation is important. In this thesis, the CV-DCM generally performs best, but its outcomes are less directly interpretable. The pixel-share MNL and the qualitative validation therefore remain important for understanding and checking the results of the CV-DCM.

6.2. Strengths and Limitations

This section discusses the strengths and limitations of the research. The strengths show how the study contributes to pedestrian route choice modelling by comparing multiple model specifications, including trip-purpose-specific models, and evaluating the models beyond standard model fit measures. The limitations show how the findings should be interpreted in relation to the survey sample, the use of SP data, and the simplified representation of walking environments through static street-level images.

6.2.1. Comparison of Multiple Model Specifications

One of the main strengths of this study is the structured comparison of three model specifications that estimate pedestrian route choice preferences using different levels of information. The basic MNL provides a benchmark based only on travel time, the pixel-share MNL adds predefined and interpretable visual attributes, and the CV-DCM uses the full street-level image. Earlier work by Sevtsuk et al. (2021) already showed that visual street characteristics can help explain pedestrian route choice, but their model only used visual attributes selected in advance. Machine learning approaches can use full images and may capture more visual information, but they are often more difficult to interpret (Van Cranenburgh et al., 2022). The comparison of these different approaches is especially relevant because the interpretability of the CV-DCM is not just treated as a limitation, but it is considered to be a driving factor of the research design. The pixel-share MNL makes it possible to interpret the role of specific visual attributes, while the CV-DCM tests whether the full image adds explanatory value beyond these predefined variables. This allows for a clearer view of the trade-offs between predictive performance and model interpretability.

6.2.2. Trip-Purpose-Specific Modelling

A second strength is the estimation of separate models for walking to public transport, walking to work or school, walking home, and walking in free time for each model specification. This makes it possible to determine how travel time and visual street-level characteristics are valued across different walking contexts, and whether similar patterns appear across all model specifications. Because context-specific differences are found in the basic MNL, pixel-share MNL, and CV-DCM, this suggests that they are less likely to be caused by the structure of one particular model and more likely to reflect patterns in the underlying choice data.

Existing literature shows that the importance of travel time, the built environment, and route attractiveness can vary across walking purposes, but trip purpose is not always modelled directly in pedestrian route choice (Hatamzadeh et al., 2014; Larranaga and Cybis, 2014; Liu et al., 2026). This thesis addresses this limitation by including trip purpose explicitly in the modelling framework. The spatial application also shows that these differences are relevant beyond a difference in estimated parameters, since it shows that the trip-purpose-specific models produce different preferred routes for the same origin and destination on the same street network.

6.2.3. Model Validation Beyond Model Fit

A third strength of this thesis is that the models are evaluated using more than estimation results and goodness-of-fit measures. As a first validation step, predictive performance is evaluated on choice tasks that were not used during model estimation through a train-test split. Several additional checks are then used to assess the reliability and behavioural meaning of the results. Model robustness is examined by comparing multiple specifications of the pixel-share MNL, testing different hyperparameter settings for the CV-DCMs, and comparing the results across different random seeds. Model consistency is tested through same-street image comparisons to determine whether images of similar street environments receive similar predicted utilities. Sensitivity is examined using AI-edited images, which show how controlled visual changes affect predicted image utility. Behavioural plausibility is analysed using perception scores and comparisons of high- and low-utility images. Finally, the spatial application examines whether the estimated utilities of the different trip-purpose models lead to plausible and meaningful differences between routes.

Parady et al. (2020) argue that choice models in transportation research are often evaluated mainly through in-sample goodness-of-fit measures, while validation beyond the estimation data is reported much less frequently. This thesis responds to this limitation by combining out-of-sample prediction with several additional forms of model evaluation. This is particularly important for the CV-DCM, because strong predictive performance does not necessarily mean that a machine learning model has learned behaviourally meaningful relationships (Van Cranenburgh et al., 2022). By combining predictive validation with robustness, consistency, sensitivity, and behavioural plausibility checks, this research provides a more critical evaluation of the models.

6.2.4. Sample Generalisability and Hypothetical Bias

A first limitation of this study concerns the generalisability of the estimated parameters and models. The survey responses were collected through a convenience sample, which means that certain population groups may be overrepresented or underrepresented. Pedestrian route choice preferences may differ according to characteristics such as age, gender, education level, mobility level, and residential environment (Basu et al., 2022; Borst et al., 2008; Mazzulla et al., 2024). This means that the estimated parameters should be interpreted within the context of the collected sample. However, the comparison between the three model specifications and trip-purpose contexts remains meaningful, because the models were estimated on the same dataset and can be compared in terms of their ability to explain the observed choices. The limitation mainly affects the extent to which the estimated preferences can be interpreted as general pedestrian preferences in the Netherlands.

A second limitation related to the data collection is the use of Stated Preference data. An SP experiment was selected because it allows route attributes and trip-purpose contexts to be varied under controlled conditions. This made it possible to analyse the trade-offs respondents made between travel time and the visual street-level characteristics. However, because the choices remain hypothetical and respondents do not experience actual walking times or physical effort, this have introduced a hypothetical bias (Jin et al., 2025). There are possibilities to combine SP and RP data to connect these controlled trade-offs to observed real-world route choices (Ben-Akiva et al., 1994). However, this would require additional data collection and introduces other methodological challenges. For this study, this limitation means that the models should be used to show which routes would be preferred, not to predict actual pedestrian routes or flows.

6.2.5. Simplification of the Walking Environment

Each route alternative in the SP experiment was represented by a single static image and a travel time. As a result, the models are estimated based on only a single visual representation of the entire route. In reality, walking is a continuous and dynamic experience in which pedestrians move through multiple streets and evaluate changing route environments along the way (Hoogendoorn and Bovy, 2004; Tong and Bode, 2022). These street environments may differ in greenery, sidewalk quality, traffic exposure, crowdedness, crossings, and overall attractiveness, all of which have been identified as relevant factors in pedestrian route choice research (Basu et al., 2022). This limitation is especially relevant for the interpretation of the Amsterdam-Zuid application. In this spatial application, multiple images were collected for each street segment and they were used to estimate the utility of individual street segments using the CV-DCM models. The image utilities were combined with the travel time utility of each segment. The utility of a complete route was then calculated by summing the utilities of the street segments along that route. This made it possible to identify the route with the highest utility for each trip-purpose model and to compare it to the objectively shortest route. However, this approach relies on the assumption that the utility of a full walking route can be represented as the sum of the utilities of its individual street segments. As this assumption was not validated in this research, the

Amsterdam-Zuid application should be interpreted as an exploratory spatial analysis. This means that the application can illustrate how visual attractiveness may vary across a network, but it should not be used as direct evidence for actual pedestrian route choices without further validation.

The final limitation is related to the assumptions of the environmental context. The images were selected to represent daytime and good weather conditions during spring and summer months. This is particularly important for interpreting the effect of greenery, because greenery was the most consistent positive visual characteristic in the model results. In the pixel-share MNL models, visible greenery had a statistically significant positive effect on route utility across all trip-purpose contexts and the qualitative interpretation of the CV-DCM results also showed that images with higher utilities often included greener street environments. However, this positive effect should not automatically be assumed to be the same across all seasons, weather conditions, or times of day. Research on environmental preferences of green spaces shows that preferences can differ between winter and summer, and they conclude that summer and winter perceptions cannot simply be considered the same (Duan and Li, 2022). Walking research focusing specifically on winter conditions also shows that this can affect route choice and walking behaviour due to slipperiness of the surface (Fossum et al., 2024). The same street environment may also be perceived differently after dark. Research on greenery and lighting shows that green areas or parks may be avoided at night when poor lighting or dense vegetation reduce visibility and perceived safety (Rahm et al., 2020). The positive effect of greenery found in this thesis is therefore context dependent, meaning that it applies specifically to the daytime, good-weather, spring and summer conditions represented in the sampled images.

6.3. Research Implications

The findings of this thesis have implications within three fields. First, they contribute to pedestrian route choice modelling by showing how visual street-level information and trip-purpose context can be included in route choice preference models. Second, they are relevant for urban planning and policy, because image-based route attractiveness can support the evaluation of walking routes beyond distance and travel time. Third, they contribute to the broader discussion on Artificial Intelligence (AI) in choice modelling and spatial policy analysis, because the CV-DCM shows both the opportunities and risks of using full-image models in behavioural and policy-oriented applications.

6.3.1. Pedestrian Route Choice Modelling

For pedestrian route choice modelling, the main implication is that the level of modelling detail should better match the purpose of the model. Shortest-path models can still be useful as a simple approximation of pedestrian route choice, but the results of this thesis show that when the aim is to understand pedestrian route preferences or evaluate the quality of walking environments, these shortest path models are not able to fully capture the complexity of pedestrian behaviour. The results show that the street environment influences how pedestrians evaluate and choose routes and that adding visual information improves predictive performance. This does not imply that every future pedestrian model should include full street-level images or a complex CV-DCM. Rather, it means that the right levels of complexity should be chosen to model pedestrian behaviour. This is especially relevant because pedestrian models are increasingly used to support urban planning and to assess how changes in the built environment can influence pedestrian behaviour, which may require different levels of complexity (Zafri and Sevtsuk, 2026).

This implication also relates to the use of trip-purpose-specific models. The same street environment may be evaluated differently depending on whether someone is walking to public transport, walking to work or school, walking home, or walking in free time. This thesis shows that the trade-offs between travel time and visual street-level characteristics differ by trip purpose within the same modelling framework, and that these differences can lead to different preferred routes. For modelling practice, this means that pedestrian route choice models should not always assume one general type of walking behaviour.

6.3.2. Route Attractiveness in Urban Planning and Policy

The spatial application in Amsterdam-Zuid showed that the routes selected by the model were often not the shortest routes. This suggests that route attractiveness can influence route preferences in the model, and that a more attractive walking environment may compensate for additional travel time. For urban planning, this means that improving the attractiveness of specific streets could make some walking routes more appealing. This idea is similar to nudging as a policy approach, where behaviour is influenced by changing the available choices and choice environment without restricting other alternatives (Steffen et al., 2024). In the context of this thesis, street design can be seen as a possible spatial form of nudging. By designing streets in visually attractive ways, planners may be able to make some routes more appealing, which could encourage pedestrians to choose these specific routes more often, although the behavioural effect would need to be validated with observed route choice data.

The CV-DCM in this thesis translates qualitative visual preferences from street-level images into numerical utility values, which makes it possible to map visual route attractiveness across the street network. With this spatial representation, it is possible to identify streets that are an important connection in the network, but are unattractive from a pedestrian perspective. For urban planning and policy, this means that the visual route choice models can complement existing tools by adding a behavioural layer of perceived environmental quality. Cities can use spatial indicators to monitor urban and transport policy, reveal inequalities within cities, and guide interventions and investments (Giles-Corti et al., 2022). The estimated route attractiveness from the CV-DCM can be considered together with other policy-relevant indicators, such as traffic safety, air pollution, or access to public transport, to better understand where urban interventions may be needed. For example, a street with low estimated attractiveness, low traffic safety and high air pollution could be a stronger candidate for improvement than a street with only one of these issues. In this way, visual models can help support the prioritisation of measures or locations and provide a clearer justification for selecting specific areas for further analysis or physical visits.

Finally, the results show that visible greenery is an important part of visual route attractiveness and pedestrian route preferences. Greenery was the most consistent positive visual street-level characteristic across the different trip-purpose contexts. While this thesis highlights the role of greenery in the walking experience, literature shows that green infrastructure can also contribute to other urban goals, such as heat mitigation, flood prevention, increased biodiversity, and urban liveability (Wang et al., 2024). This means that adding greenery to improve the attractiveness of walking routes may also support broader environmental and liveability goals. While adding greenery should not be seen as a universal solution, this thesis shows that greenery also supports pedestrian route attractiveness. Greenery may therefore help connect walking policy with broader liveability and environmental goals.

6.3.3. Artificial Intelligence in Behavioural Modelling and Policy Analysis

This thesis is also part of a broader shift in which AI, machine learning, and CV are becoming more relevant for behavioural modelling and policy analysis. Traditional choice models often rely on variables that are structured and easy and describe, such as travel time, cost, or categorical attributes. However, many real-world choices are also influenced by information that is harder to describe in a numerical way, such as the visual appearance of a street. Machine learning and CV make it possible to use richer data sources, such as images, text, and video (Van Cranenburgh et al., 2022). In this thesis, the CV-DCM uses full street-level images directly in the utility function, which allows the model to capture visual information that is not fully represented by predefined variables. Research on visual information in stated choice experiments also shows that images can make respondent choices more consistent and increase confidence (Kabaya et al., 2024). This suggests that adding visual information can help respondents better understand the choice situation, while also allowing CV models to capture visual aspects that influence these decisions. At the same time, route choice behaviour remains influenced by habits, personal preferences, context, uncertainty, and other unobserved factors (Basu et al., 2022; Duives, 2025). This means that CV can improve the estimation of stated route preferences by capturing more visual information, but it does not remove the uncertainty of human behaviour.

A second implication is that the CV-DCM can make walkability analysis both more local and more scalable. Literature on AI in urban planning shows that it has potential because it can process large and complex urban data and support data-driven decision-making (Lartey and Law, 2025). By training a CV-DCM on survey or image data from a specific city, neighbourhood, or population group, the estimated preferences can better reflect local contexts. At the same time, the CV component can analyse large numbers of street-level images relatively quickly, making it possible to create a spatial overview of route attractiveness across a larger area. This could help policymakers compare streets or neighbourhoods and explore how visual characteristics, such as greenery or traffic exposure, may affect perceived route attractiveness. In this way, these models could become useful as part of data-driven policy analysis and urban planning, especially when they are combined with existing spatial data, local knowledge, and expert knowledge.

However, the use of AI in behavioural modelling should be approached with caution, as it makes the role of the modeller even more important. A model that can quickly interpret and process thousands of images may seem objective, but its results still depend on the data, modelling choices, and the way the outputs are interpreted. Studies on the ethical use of AI in urban planning show that these tools can create risks related to transparency, accountability, bias, privacy, fairness, and public involvement (Sanchez et al., 2024; Lartey and Law, 2025). This is relevant for the CV-DCM, because the model can estimate whether an image is associated with higher or lower utility, but the image features learned by the model do not have a direct behavioural interpretation (van Cranenburgh and Garrido-Valenzuela, 2025). This means that it is not immediately clear which visual elements drive the estimated image utility. The results also show that the performance of the CV-DCM is sensitive to modelling choices, such as the selected hyperparameter configuration. The models also respond unpredictably to unexpected visual elements, as shown by the addition of visible trash, because they can only learn from patterns that are present in the training data. Therefore, the CV-DCM models could be useful for exploring visual route preference patterns, but they are not yet suitable as stand-alone policy instruments. Before these models can be used in decision-making, their route preferences should be validated with real route choice data, compared with more interpretable models, and, where possible, supported by Explainable Artificial Intelligence (XAI) methods.

Finally, caution is also needed when using AI-generated or AI-edited images. In this thesis, edited images were useful for controlled sensitivity analysis, because they made it possible to test how the model responds to specific visual changes. However, these images are not representative of real interventions and should not be treated as direct evidence of how actual street redesigns would affect behaviour. They can support exploration, but they may also create misleading confidence if they are presented without enough context. Calzada and Eizaguirre (2025) argue that AI in urban planning should be connected to local context, public values, and the people affected by planning decisions, rather than being treated as only a technical solution. This means that generated or edited images can be used as scenario tools, but not as proof of policy effects. They can help explore how model predictions respond to visual changes, but human judgement remains necessary before the results are used in real planning decisions.

6.4. Recommendations and Future Research

Based on the findings, limitations, and implications of this thesis, this section presents recommendations for both practical applications and scientific research. The practical recommendations focus on how visual route choice models can be used responsibly in pedestrian modelling and urban policy analysis. The scientific recommendations focus on how future research could further improve visually enriched pedestrian route choice models, for example by using better data, improving the modelling approach, making the results easier to interpret, and testing the models in other contexts.

6.4.1. Practical Recommendations

The first recommendation is to use visual route choice models only when they match the purpose of the analysis. If the aim is to make a simple approximation of pedestrian movement, a shortest-path or travel-time-based model may be sufficient. However, if the aim is to evaluate walkability, route attractiveness, or street-level interventions, visual street-level quality should be included. The level of model complexity should therefore depend on the policy or modelling question.

A second recommendation is to use the CV-DCM mainly as an exploratory decision-support tool. For applied research organisations such as TNO, the method can be applied or further developed for specific case study areas, such as Amsterdam-Zuid. The model can be used to compare streets or neighbourhoods, identify locations of interest, and support the prioritisation of possible interventions. However, the model estimated in this thesis should not be used directly to predict real pedestrian flows, because it is based on stated preferences and has not yet been validated with observed route choices.

A third recommendation is to estimate local models when the method is used to support local urban policy. If the method is applied in a specific city or neighbourhood, it can be useful to collect data from people who live, work, study, or regularly walk in that area. This would make the estimated preferences more connected to the local population and would reduce the risk of applying preferences from one context to another. This is important because pedestrian preferences can differ between places, trip purposes, population groups, and walking contexts (Basu et al., 2022; Duives, 2025; Levine and Norenzayan, 1999).

A final recommendation is to communicate model uncertainty clearly. The CV-DCM can capture visual information from full street-level images, but the learned image features do not have a direct behavioural interpretation (van Cranenburgh and Garrido-Valenzuela, 2025). Model outputs should therefore be validated, compared with more interpretable models, and used to support discussion and prioritisation. They should not be used to make planning decisions automatically, especially because AI-based tools in planning can raise concerns about transparency, accountability, bias, privacy, fairness, and lack of public involvement (Sanchez et al., 2024; Lartey and Law, 2025).

6.4.2. Future Research

A first direction for future research is to use data collection methods that are closer to actual walking behaviour and that represent walking routes more realistically. This thesis used an SP experiment in which each route alternative was represented by a travel time and one street-level image. This made it possible to study controlled trade-offs between travel time and visual street-level information, but it also simplified the walking experience. Future research could combine Stated Preference (SP) and Revealed Preference (RP) data to examine to what extent the visual preferences found in stated choice tasks also influence actual route choices. RP data, such as GPS trajectories, pedestrian counts, or observed route choices, could help connect stated preferences to real walking behaviour. At the same time, other visualisation methods such as multiple images per route, image sequences, videos or virtual reality could be used to represent walking routes more realistically while still maintaining experimental control (Arellana et al., 2020).

A second direction is to design experiments that focus more directly on specific visual attributes and possible street-level interventions. In this thesis, the images were selected across the ten largest cities in the Netherlands to represent a range of combinations of greenery, traffic, and people presence. Future experiments could use more controlled image sets to isolate the effect of specific visual characteristics. These could include the attributes used in this thesis, such as greenery, traffic exposure, and people presence, but also other factors such as sidewalk quality, litter, building quality, lighting, crossings, or pedestrian-specific amenities. This would make it possible to examine how specific visual elements influence pedestrian route preferences and which types of interventions are most likely to affect route choice preferences. This kind of research could also make the link between route choice modelling and urban design more direct.

A third direction is to further examine how visual route preferences differ across trip purposes, population groups, and spatial contexts. This thesis showed that travel time and visual street-level information are valued differently across walking to public transport, walking to work or school, walking home, and walking in free time. Future research could examine these differences in more detail by adding more trip-purpose contexts or by using DCM specifications that account for preference heterogeneity, such as mixed logit or latent class models. This is also relevant for differences between groups of pedestrians, since preferences may vary by age, gender, walking habits, familiarity with the area, residential environment, or mobility limitations (Basu et al., 2022; Duives, 2025). In addition, future studies should test whether similar patterns are found in other cities, neighbourhoods, seasons, and weather conditions. This would help determine whether the estimated visual preferences are generalisable or mainly context-specific.

Future research should also focus on making the CV-DCM more transparent and easier to explain. The CV-DCM can capture visual information from full street-level images, but its learned image representation is not directly interpretable. This limits its usefulness for policy and planning, because decision-makers need to understand why certain streets or routes receive higher or lower utility before using model outcomes to justify interventions. Future research could develop hybrid models that combine predefined visual attributes with learned image features. This would make it possible to benefit from the predictive strength of the CV-DCM, while keeping part of the model behaviourally interpretable. In addition, Explainable Artificial Intelligence (XAI) methods could be tested to examine which visual elements the model responds to, and whether these elements are behaviourally meaningful. Improving explainability would make image-based route choice models more useful as decision-support tools in urban policy and planning.

7

Conclusion

This thesis examined how visual and non-visual route attributes influence pedestrian route choice preferences in the Netherlands. By combining a Stated Preference (SP) experiment with street-level images and discrete choice modelling, the research investigated whether visual information can improve the prediction and interpretation of stated pedestrian route choice preferences, and how these preferences differ across walking contexts. This chapter concludes the thesis by answering the research questions, presenting the main conclusion, and summarising the contributions of the research.

7.1. Answer to the Research Questions

This section answers the sub-questions of this research and uses these answers to address the main research question:

To what extent do visual and non-visual route attributes influence pedestrian route choice preferences in the Netherlands?

1. Which visual and non-visual route attributes are relevant to pedestrians when choosing a route?

Pedestrian route choice is influenced by a combination of trip efficiency, built- and natural-environment characteristics, socio-demographic factors, and contextual factors. In existing research, travel time or distance is often identified as one of the most important attributes in pedestrian route choice. However, pedestrians also respond to visual characteristics of the walking environment, such as the amount of greenery, the availability of pedestrian space, the presence of traffic, and the level of crowdedness. These attributes can affect how comfortable, safe, or attractive a route is perceived to be. In addition to these physical and visual attributes, contextual factors such as sound, weather, smell, familiarity with the area, and the purpose of the walking trip may also influence route choice. Since not all potentially relevant attributes could be examined within the scope of this thesis, this research focused on a selected set of attributes: travel time, visible greenery, sidewalk availability, visible traffic, crowdedness, and the trip-purpose context in which the route choice was made: walking to public transport, walking to work or school, walking home, and walking in free time.

2. How can the trade-offs pedestrians make between visual and non-visual route attributes be measured?

The trade-offs between visual and non-visual route attributes were measured using a Stated Preference (SP) experiment in which respondents chose between simplified route alternatives. This approach makes it possible to systematically vary the attributes of the alternatives, so that the effect of changes in individual attributes can be identified. In this thesis, the experiment was implemented as an online survey in which respondents repeatedly chose between two route alternatives that differed in travel time and visual street-level information shown through images. This was done across different trip-purpose contexts. By using an efficient experimental design to maximise the information obtained from these choices, the SP experiment produced a dataset that could be used to estimate how pedestrians trade off travel time against visual characteristics of the walking environment.

3. How can visual and non-visual route attributes be incorporated into a discrete choice model for pedestrian route choice preferences?

Visual and non-visual route attributes can be incorporated into a Discrete Choice Model (DCM) in different ways. This thesis estimated three types of models to compare how the inclusion of visual information affects the explanation of pedestrian route choice preferences. The first model is a basic Multinomial Logit (MNL) model that only includes travel time as an attribute. This model is used as a baseline to assess the added explanatory value of visual information. The second model is a pixel-share MNL model, in which travel time is combined linearly with predefined pixel-share variables obtained through semantic segmentation. These variables represent the selected visual attributes as the share of visible greenery, pedestrian infrastructure, traffic, and people in the image. This model incorporates visual information in an explicit and interpretable way. The third model is the Computer Vision-Enriched Discrete Choice Model (CV-DCM), in which travel time is combined with the full street-level image in the utility specification. This allows the model to learn relevant visual patterns from the complete image, rather than relying only on predefined visual variables. However, this also means that the visual component of the model is less directly interpretable, as it does not produce parameter estimates for individual visual attributes.

4. How do pedestrian route choice preferences and model performance change when visual attributes are included, and how do these effects differ across trip-purpose contexts?

Including visual attributes improves predictive performance on the collected data. Both the pixel-share Multinomial Logit (MNL) model and the Computer Vision-Enriched Discrete Choice Model (CV-DCM) generally outperformed the basic MNL model, with the CV-DCM producing the best predictive results on the collected data. This indicates that travel time alone does not sufficiently explain the choices observed in the experiment, and that visual information adds explanatory value. In the pixel-share MNL, the amount of visible greenery showed the clearest and most consistent positive association with route preferences across contexts. The qualitative validation of the CV-DCM also suggested that images with higher predicted utility were often greener and more attractive. However, the general effect of visual information differed across trip-purpose contexts. For more goal-oriented trips, such as walking to public transport or walking to work or school, travel time remained important, although visual information still contributed to route utility. For less time-constrained trips, especially walking in free time, visual attributes explained a larger share of the observed choice variation, while the relative importance of travel time decreased. The results show that pedestrian route choice preferences are not explained by travel time alone, and that the relative importance of travel time and visual street-level attributes depends on the purpose of the walking trip.

5. How can a visually-enriched route choice model be applied to analyse trip-purpose-specific route preferences across a pedestrian network?

A trip-purpose-specific Computer Vision-Enriched Discrete Choice Model (CV-DCM) can be applied to a pedestrian network by estimating route utilities between an origin and destination and comparing the predicted preferred routes with the objectively shortest route. In this thesis, this was demonstrated in an Amsterdam-Zuid case study. The application showed that the preferred routes were not always the shortest routes and that these routes differed between trip-purpose models. This indicates that the same streets can receive different estimated utilities depending on the walking context. However, these results should be interpreted as exploratory rather than as direct predictions of actual pedestrian behaviour in Amsterdam-Zuid, as the application mainly demonstrates how the CV-DCM can be applied to a pedestrian network and used to analyse spatial differences in route attractiveness.

Answer to the Main Research Question

The results show that visual and non-visual route attributes both influence pedestrian route choice preferences in the Dutch urban context of this thesis. Travel time remains an important attribute, especially for more goal-oriented trips such as walking to public transport or walking to work or school. However, travel time alone does not sufficiently explain the choices observed in the Stated Preference (SP) experiment. Visual street-level attributes add explanatory value, as models that included visual information performed better at explaining observed choices than the model based only on travel time. In the pixel-share MNL, visible greenery showed the clearest and most consistent positive association with route choice preferences. The qualitative validation of the CV-DCM also suggested that images with higher predicted utility were often greener and more attractive. The influence of visual information was strongest for less time-constrained trips, especially walking in free time, where visual attributes explained a larger share of the observed choice variation. Therefore, within the collected SP dataset, visual and non-visual route attributes both influence pedestrian route choice preferences to a meaningful extent. Their relative importance depends on the context in which the walking trip takes place.

7.2. Main Conclusion

The main conclusion of this thesis is that the pedestrian route choice preferences in the data cannot be fully represented by travel time alone. This was shown by comparing a basic Multinomial Logit (MNL) model using only travel time with two models that included visual information: a pixel-share MNL model using predefined visual attributes and a Computer Vision-Enriched Discrete Choice Model (CV-DCM) using full street-level images. Both image-based models generally outperformed the basic MNL model, with the CV-DCM producing the strongest predictive results in most contexts. This shows that the visual quality of the walking environment matters for the stated preferences observed in this study. The results also show that route choice preferences differ across trip-purpose contexts. Travel time is more important for goal-oriented trips, while visual street-level characteristics become more important for less time-constrained trips. This means that pedestrian route choice preferences should not always be represented with one general preference structure across all walking contexts. Instead, the purpose of the walking trip affects how pedestrians trade off travel time against the quality of the walking environment.

At the same time, the results should be interpreted within the scope of the research design. The Stated Preference (SP) experiment made it possible to measure controlled trade-offs between travel time and visual street-level information, but the findings are based on stated choices from a specific sample rather than observed walking behaviour from a representative population. Therefore, the results should not be interpreted as direct predictions of all pedestrian route choices in the Netherlands. Instead, they show that visual information can meaningfully enrich pedestrian route choice modelling and that trip-purpose-specific preferences are important to consider.

7.3. Contributions

Based on the methodology and results of this thesis, three main contributions can be highlighted. First, to the best of the author's knowledge, this thesis is the first to apply the CV-DCM to pedestrian route choice modelling. This extends the use of image-based discrete choice modelling to a pedestrian context and shows how full street-level images can be used to explain pedestrian route choice preferences.

Second, this thesis provides a comparison of three models that differ in complexity and in how visual information is included: a basic Multinomial Logit (MNL) model based only on travel time, a pixel-share MNL model using predefined visual attributes, and a CV-DCM that learns from full street-level images. This comparison shows the added value of including visual information when explaining stated route choice preferences. It also highlights the trade-off between interpretability and predictive performance. The pixel-share models are easier to interpret, while the CV-DCM can capture broader visual patterns in the street-level images.

Third, this thesis contributes to the link between walkability research and route choice modelling. Walkability research often focuses on the quality of the walking environment, while route choice modelling often focuses on route efficiency. This thesis shows that both perspectives are relevant for pedestrian route choice. The results show that, within the stated choice experiment, pedestrian route preferences are both visually and context dependent. In particular, the relative importance of travel time and visual street-level characteristics differs across trip-purpose contexts. These differences are visible in the model results, the qualitative image comparisons, and the spatial route application.

Overall, this thesis shows that visual street-level information can be incorporated into pedestrian route choice models at different levels of complexity. The findings therefore support a broader view of pedestrian route choice, in which walking routes are not only evaluated as connections between origins and destinations, but also as environments that are experienced while walking.

Bibliography

Agrawal, A., Schlossberg, M., and Irvin, K. (2008). How far, by which route and why? a spatial analysis of pedestrian preference. *Journal of Urban Design*, 13:81–98.

Arellana, J., Garzón, L., Estrada, J., and Cantillo, V. (2020). On the use of virtual immersive reality for discrete choice experiments to modelling pedestrian behaviour. *Journal of Choice Modelling*, 37:100251.

Basu, N., Haque, M. M., King, M., Kamruzzaman, M., and Oviedo-Trespalacios, O. (2022). A systematic review of the factors associated with pedestrian route choice. *Transport Reviews*, 42(5):672–694.

Basu, N., Oviedo-Trespalacios, O., King, M., Kamruzzaman, M., and Haque, M. M. (2023). What do pedestrians consider when choosing a route? the role of safety, security, and attractiveness perceptions and the built environment during day and night walking. *Cities*, 143:104551.

Batty, M. (2009). Urban modeling. In Kitchin, R. and Thrift, N., editors, *International Encyclopedia of Human Geography*, pages 51–58. Elsevier, Oxford.

Bekhor, S., Ben-Akiva, M. E., and Ramming, M. S. (2006). Evaluation of choice set generation algorithms for route choice models. *Annals of Operations Research*, 144(1):235–247.

Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H., and Rao, V. (1994). Combining revealed and stated preferences data. *Marketing Letters*, 5(4):335–349.

Ben-Akiva, M. E. and Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, Cambridge, MA.

Bliemer, M. and Rose, J. (2005). Efficiency and sample size requirements for stated choice studies.

Bliemer, M. and Rose, J. (2024). *Designing and conducting stated choice experiments*, pages 172–205.

Bliemer, M. C. and Collins, A. T. (2016). On determining priors for the generation of efficient stated choice experimental designs. *Journal of Choice Modelling*, 21:10–14. Standalone technical contributions in choice modelling.

Borchers, T., Wittowsky, D., and Fernandes, R. (2024). A comprehensive survey and future directions on optimising sustainable urban mobility. *IEEE Access*, PP:1–1.

Borst, H. C., Miedema, H. M., de Vries, S. I., Graham, J. M., and van Dongen, J. E. (2008). Relationships between street characteristics and perceived attractiveness for walking reported by elderly people. *Journal of Environmental Psychology*, 28(4):353–361.

Bovy, P. H. L. and Stern, E. (1990). *Route choice: Wayfinding in transport networks*.

Brownstone, D., Bunch, D. S., and Train, K. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research Part B: Methodological*, 34(5):315–338.

Calzada, I. and Eizaguirre, I. (2025). Digital inclusion and Urban AI: strategic roadmapping and policy challenges. *Discover Cities*, 2(1).

Centraal Bureau voor de Statistiek (2024). Hoeveel reisden inwoners van Nederland en hoe?

Cerin, E., Leslie, E., du Toit, L., Owen, N., and Frank, L. D. (2007). Destinations that matter: Associations with walking for transport. *Health Place*, 13(3):713–724.

Cheng, B., Schwing, A. G., and Kirillov, A. (2022). Masked-attention mask transformer for universal image segmentation. *arXiv preprint arXiv:2112.01527*.

Choi, J. and Kang, Y. (2025). Modeling and explaining perceived walkability in urban environments using street view images and explainable AI. *Spatial Information Research*, 33(6).

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Li, F.-F. (2009). Imagenet: a large-scale hierarchical image database. pages 248–255.

Duan, Y. and Li, S. (2022). Study of different vegetation types in green space landscape preference: Comparison of environmental perception in winter and summer. *Sustainability*, 14(7).

Dubey, A., Naik, N., Parikh, D., and Raskar, R. (2016). Deep learning the city : Quantifying urban perception at a global scale. volume 9905.

Duives, D. C. (2025). Pedestrian travel choice behavior and modeling. In Daamen, W. and Duives, D. C., editors, *Walking and Pedestrians*, volume 15 of *Advances in Transport Policy and Planning*. Elsevier.

European Commission (2021). The new eu urban mobility framework. Technical Report COM(2021) 811 final, European Commission, Brussels.

Fernández Aguilar, C., Brosed-Lázaro, M., and Carmona-Derqui, D. (2023). Effectiveness of mobility and urban sustainability measures in improving citizen health: A scoping review. *International Journal of Environmental Research and Public Health*, 20:2649.

Fosgerau, M., Frejinger, E., and Karlstrom, A. (2013). A link based network route choice model with unrestricted choice set. *Transportation Research Part B: Methodological*, 56:70–80.

Fossum, M., Hillnhütter, H., and Ryeng, E. O. (2024). Winter walking – the effect of winter conditions on pedestrians' step length and step frequency. *Transportmetrica A: Transport Science*, 20(1):2122760.

Garrido-Valenzuela, F., Cats, O., and van Cranenburgh, S. (2025). Pixelsurvey: An open-source framework for creating surveys with yaml.

Giles-Corti, B., Moudon, A., Lowe, M., Cerin, E., Boeing, G., Frumkin, H., Salvo, D., Foster, S., Kleeman, A., Bekessy, S., Sa, T., Nieuwenhuijsen, M., Higgs, C., Hinckson, E., Adlakha, D., Arundel, J., Liu, S., Oyeyemi, A., Nitvimol, K., and Sallis, J. (2022). What next? expanding our view of city planning and global health, and implementing and monitoring evidence-informed policy. *The Lancet Global Health*, 10:e919–e926.

Golledge, R. G. (1992). Place recognition and wayfinding: Making sense of space. *Geoforum*, 23(2):199–214.

Google (2026). Google street view. `https://www.google.com/streetview/`.

Hatamzadeh, Y., Habibian, M., and Khodaii, A. (2014). Walking behaviors by trip purposes. *Transportation Research Record: Journal of the Transportation Research Board*, 2464:118–125.

Hoogendoorn, S. and Bovy, P. (2004). Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological*, 38(2):169–190.

Jin, C.-J., Li, J., Wu, C., Xiu, X., and Li, D. (2025). Findings in pedestrian route choice experiments: Tradeoff and hypothetical bias. *Travel Behaviour and Society*, 40:101000.

Kabaya, K., Tajima, K., Ichinose, D., and Asano, M. (2024). Do different visual presentation formats encourage different choice behaviors? discrete choice experiment on urban park landscapes. *Environmental Economics and Policy Studies*, 27.

Kennisinstituut voor Mobiliteitsbeleid (2024). Loopfeiten 2024.

Kouwenhoven, M., Muller, J., Thoen, S., Willigers, J., and de Jong, G. (2023). Values of time, reliability and comfort in the netherlands 2022: New values for passenger travel and freight transport. Technical report, Significance. Commissioned by KiM Netherlands Institute for Transport Policy Analysis.

Larranaga, A. and Cybis, H. (2014). The relationship between built environment and walking for different trip purposes in porto alegre, brazil. *International Journal of Sustainable Development and Planning*, 9:568–580.

Lartey, D. and Law, K. M. (2025). Artificial intelligence adoption in urban planning governance: A systematic review of advancements in decision-making, and policy making. *Landscape and Urban Planning*, 258:105337.

Levine, R. and Norenzayan, A. (1999). The pace of life in 31 countries. *Journal of Cross-Cultural Psychology*, 30.

Lin, H., Sun, G., and Li, R. (2015). The influence of built environment on walking behavior: Measurement issues, theoretical considerations, modeling methodologies and chinese empirical studies. *Space-Time Integration in Geography and GIScience: Research Frontiers in the US and China*, pages 53–75.

Liu, X., Wong, S., Li, B., Chng, S., and Neo, H. (2026). Walking the preferred path: exploring spatial-temporal variation in pedestrian route choice through revealed preferences in a developing city. *Travel Behaviour and Society*, 44:101278.

López-Lambas, M. E., Sánchez, J. M., and Alonso, A. (2021). The walking health: A route choice model to analyze the street factors enhancing active mobility. *Journal of Transport Health*, 22:101133.

Ma, X., Ma, C., Wu, C., Xi, Y., Yang, R., Peng, N., Zhang, C., and Ren, F. (2021). Measuring human perceptions of streetscapes to better inform urban renewal: A perspective of scene semantic parsing. *Cities*, 110:103086.

Mapillary (2026). Mapillary street-level imagery. `https://www.mapillary.com`.

Mazzulla, G., Eboli, L., and Forciniti, C. (2024). Do women perceive pedestrian path attractiveness differently from men? *Transportation Research Part A: Policy and Practice*, 179:103890.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, 3(4):303–328.

Ministry of Infrastructure and Water Management (2022). Mobility vision 2050: Framework memorandum.

Mondschein, A. (2018). Persistent differences in walking across the socioeconomic spectrum: Variations by trip purpose. *Journal of Planning Education and Research*, 41:0739456X1879665.

Neuhold, G., Ollmann, T., Rota Bulò, S., and Kontschieder, P. (2017). The mapillary vistas dataset for semantic understanding of street scenes. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.

OpenStreetMap contributors (2017). Planet dump retrieved from https://planet.osm.org . `https://www.openstreetmap.org`.

Parady, G., Ory, D., and Walker, J. (2020). The overreliance on statistical goodness-of-fit and under-reliance on model validation in discrete choice models: A review of validation practices in the transportation academic literature. *Journal of Choice Modelling*, 38:100257.

Prashker, J. N. and Bekhor, S. (2004). Route choice models used in the stochastic user equilibrium problem: A review. *Transport Reviews*, 24(4):437–463.

Prato, C. G. (2009). Route choice modeling: past, present and future research directions. *Journal of Choice Modelling*, 2(1):65–100.

Rahm, J., Sternudd, C., and Johansson, M. (2020). “In the evening, I don’t walk in the park”: The interplay between street lighting and greenery in perceived safety. *URBAN DESIGN International*, 26(1):42–52.

Rose, J. and Bliemer, M. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews*, 29.

Rossetti, T., Lobel, H., Rocco, V., and Hurtubia, R. (2019). Explaining subjective perceptions of public spaces as a function of the built environment: A massive data approach. *Landscape and Urban Planning*, 181:169–178.

Sanchez, T. W., Brenman, M., and Ye, X. (2024). The Ethical Concerns of Artificial Intelligence in Urban Planning. *Journal of the American Planning Association*, 91(2):294–307.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):15–18.

Sevtsuk, A., Basu, R., Li, X., and Kalvo, R. (2021). A big data approach to understanding pedestrian route choice preferences: Evidence from san francisco. *Travel Behaviour and Society*, 25:41–51.

Steffen, J., Hook, H., and Witlox, F. (2024). Improving interest in public, active, and shared travel modes through nudging interventions. *Transportation Research Part F: Traffic Psychology and Behaviour*, 103:353–367.

Terra, R. W., Garrido-Valenzuela, F., Cats, O., and Van Cranenburgh, S. (2025). Understanding cycling route choice behaviour through street-level images and discrete choice models. *SSRN Electronic Journal*.

Ton, D., Duives, D. C., Cats, O., Hoogendoorn-Lanser, S., and Hoogendoorn, S. P. (2019). Cycling or walking? determinants of mode choice in the netherlands. *Transportation Research Part A: Policy and Practice*, 123:7–23.

Tong, Y. and Bode, N. W. F. (2022). The principles of pedestrian route choice. *Journal of The Royal Society Interface*, 19(189):20220061.

Train, K. (2009). *Discrete Choice Methods With Simulation*, volume 2009.

Ultralytics (2023). Ultralytics yolov8. `https://github.com/ultralytics/ultralytics`.

van Cranenburgh, S. and Garrido-Valenzuela, F. (2025). Computer vision-enriched discrete choice models, with an application to residential location choice. *Transportation Research Part A: Policy and Practice*, 192:104300.

Van Cranenburgh, S., Wang, S., Vij, A., Pereira, F., and Walker, J. (2022). Choice modelling in the age of machine learning - Discussion paper. *Journal of Choice Modelling*, 42:100340.

Wang, D., Xu, P.-Y., An, B.-W., and Guo, Q.-P. (2024). Urban green infrastructure: bridging biodiversity conservation and sustainable urban development through adaptive management approach. *Frontiers in Ecology and Evolution*, 12.

Watson, K., Whitfield, G., Bricka, S., and Carlson, S. (2021). Purpose-based walking trips by duration, distance, and select characteristics, 2017 national household travel survey. *Journal of Physical Activity and Health*, 18:S86–S93.

Yang, Y. and Diez Roux, A. (2012). Walking distance by trip purpose and population subgroups. *American journal of preventive medicine*, 43:11–9.

Zafri, N. M. and Sevtsuk, A. (2026). Advancing Pedestrian Models: A Comparative Review and Vision for the Future. *Journal of the American Planning Association*, pages 1–18.

Zhu, W. and Si, W. (2024). Predicting choices of street-view images: A comparison between discrete choice models and machine learning models. *Journal of Choice Modelling*, 50:100470.

A

Factors Influencing Pedestrian Route Choice

Table A provides an overview of the factors associated with pedestrian route choice identified by Basu et al. (2022). Following the grouping used in the review, the factors are organised into trip characteristics, built- and natural-environment factors, and socio-demographic factors.

Main Category	Sub-category	Specific factor	Direction
Trip Characteristics	Distance / route length	Trip length / distance of the route	Mixed
		Length of street with department stores / restaurants	Positive
		Length of street without department stores / restaurants	Positive
		Shortest path / route	Positive
	Speed and traffic	Average speed	Positive
		Maximum speed	Positive
		Speed limit	Mixed
	Traffic volume	Average daily traffic / traffic volume	Mixed
	Walking time	Travel / walking time	Mixed
		Waiting time	Mixed
	Route efficiency	Quickest path	Positive
Built- and Natural-Environment Factors	Sidewalk characteristics	% of sidewalk along route	Positive
		Sidewalk width	Mixed
		Sidewalk condition	Positive
		Sidewalk continuity	Mixed
	Street crossing facilities	Crosswalk	Mixed
		Intersection density	Mixed
		Junctions (per km)	Positive
		Mid-block crossing	Unclear
		Number of intersections / road crossings along route	Negative
		Number of lanes to cross at the end of a street segment	Positive
		Un-signalized arterial crossing	Negative
		Unmarked collector crossings	Negative
		Zebra crossings	Mixed
		Yellow pedestrian sign	Negative
		Pedestrian signals	Positive
		Traffic light / traffic signal	Mixed

Continued on next page

Main Category	Sub-category	Specific factor	Direction
Built- and Natural-Environment Factors	Route infrastructure / pedestrian network infrastructure	Nodes	Positive
		Turns	Mixed
		Street lights	Positive
		Street signs	Mixed
		Stop signs	Mixed
		Speed bumps	Negative
		Medians	Negative
		Paved surface / paved or planted median	Mixed
	Pedestrian amenities and urban design features along sidewalk	Garbage bin / litter on street	Positive
		Graffiti	Positive
		Trees / tree shades / green strips / vegetation	Mixed
		Presence of open space / park	Mixed
		Pedestrian friendly parcels (PFPs) density	Positive
		Retail presence / frontage (such as coffee shops)	Mixed
		Front gardens	Positive
		Blind walls	Mixed
		Sculpture / mural / public art / street art	Mixed
		Squares	Positive
		Bushes	Mixed
		Flowers	Positive
		Bodies of water	Positive
		Fountains	Positive
		Benches	Positive
		Shade / weather protection	Positive
		Pavement cleanliness	Positive
		Vending machines	Positive
		Bus stops	Positive
	Land use along the route	Land-use mix along route	Mixed
		Residential land-use along route	Mixed
		Commercial land-use / small commercial (shops) buildings along route	Mixed
		Institutional land uses (offices) along route	Positive
		Industrial land uses along route	Mixed
		Recreational land uses along route	Mixed
		Vacant land along route	Negative
		Traffic area along route	Negative
	Building types along the route	Tall residential buildings along route	Mixed
		Modern building type along route	Positive
		Old building type along route	Positive
		Average building height along route	Positive
		Abandoned buildings along route	Positive
	Building setback	Building setback along the route	Positive
	Density	Pedestrian density	Mixed
		Residential density	Mixed
		Non-residential density	Positive
	Safety	Traffic safety	Mixed
		Safe crossing places	Positive
		Traffic crash risk	Negative
		Interaction with motorized traffic	Negative
		Obstacles	Negative
		Detour	Negative

Continued on next page

Main Category	Sub-category	Specific factor	Direction
Built- and Natural-Environment Factors	Security	Personal security (crime safety, social environment)	Mixed
		Presence of unattended dog	Negative
	Quality of the walking environment	Accessibility	Mixed
		Familiarity	Positive
		Attractiveness	Positive
		Convenience	Positive
		Functionality	Mixed
		Crowdedness	Mixed
		Aesthetics / good scenery	Positive
		Route directness	Positive
		Connectivity	Mixed
		Path size (measure of route independence among route alternatives)	Mixed
		Comfort	Positive
		Construction noise	Negative
		Pollution	Negative
		Busy street	Negative
		Substandard street	Negative
		BE condition	Negative
		Straight path	Positive
		Diagonal path	Negative
	Topography	Terrain / steep upslope	Mixed
		Stairs	Negative
Socio-demographic Factors	Demographics	Age	Mixed
		Gender	Mixed
		Income	Positive
		Employment status / occupation	Unclear
		Race / ethnicity	Unclear
	Travel company	Company / travelling together	Positive

B

Street-Level Image Extraction and Processing

This appendix chapter provides the pseudocode and details of the street-level image extraction and processing pipeline. This pipeline consists of four main steps:

1. Extraction of pedestrian networks in Section B.1.
2. Sampling and collection of street-level images in Section B.2.
3. Image analysis and classification in Section B.3.
4. Construction of visual attributes in Section B.4.

B.1. Step 1: Pedestrian Network Extraction

Algorithm 1 Pedestrian Network Extraction

```
Input: List of cities, OpenStreetMap data
for each city do
    Download street network using OSMnx
    Filter network to contain only pedestrian-accessible segments
    Remove non-walkable and restricted segments
    Simplify network geometry
    Convert network into nodes and edges
    Save network as GeoPackage
end for
Output: Pedestrian network per city
```

OpenStreetMap Pedestrian Roads Filter

```
custom_filter = (
    '["highway"~"^(footway|path|pedestrian|living_street|residential|service|track|steps|unclassified|tertiary)$"]'
    '["area"!="yes"]'
    '["access"!~"^(no|private)$"]'
    '["foot"!~"^(no|private)$"]'
    '["motorroad"!="yes"]'
)
```

Listing B.1: Filter used to find pedestrian accesible roads in OpenStreetMap

B.2. Step 2: Street-Level Image Sampling and Collection

Algorithm 2 Street-Level Image Sampling and Collection

```
Input: Pedestrian network per city, optional city centre reference point, Street View API
for each city do
    Load pedestrian network
    Compute edge sampling probabilities based on edge length
    if centre weighting is used then
        Increase sampling probability for edges closer to the city centre
    end if
    while required number of image groups not reached do
        Sample street segment and random point based on sampling probability
        if point too close to previous samples then
            continue
        end if
        Generate nearby points along the segment
        for each point do
            Compute heading and request Street View metadata
            if no imagery or duplicate panorama then
                continue
            end if
            Check capture month and image quality (brightness, tilt)
            if requirements not met then
                continue
            end if
            Save image metadata
        end for
        if at least one image is accepted then
            Accept image group
        end if
    end while
end for
Output: Street-level image dataset
```

City Centre Reference Locations

City	Location	Latitude	Longitude
Amsterdam	De Dam	52.3739	4.8922
Rotterdam	Markthal	51.9200	4.4875
Den Haag	Grote Markt	52.0758	4.3093
Utrecht	Domtoren	52.0909	5.1213
Eindhoven	18 Septemberplein	51.4415	5.4792
Groningen	Grote Markt	53.2190	6.5678
Almere	Stadhuisplein	52.3719	5.2213
Tilburg	Heuvel	51.5577	5.0904
Breda	Grote Markt	51.5887	4.7762
Nijmegen	Grote Markt	51.8475	5.8637

Table B.1: City centre reference locations and coordinates used for each city

B.3. Step 3: Image Analysis and Filtering

Algorithm 3 Image Analysis and Filtering

```
Input: Image metadata, Street View API
for each image do
    Apply semantic segmentation model
    Assign each pixel to a visual class
    Apply object detection model
    Count pedestrians and vehicles
    Compute spatial features: road share, sidewalk share, greenery share, etc.
    Label image using rule-based criteria
    if image represents pedestrian-accessible street then
        Keep image
    else
        Reject image
    end if
    Store features and labels
end for
Output: Filtered image dataset with extracted features
```

Image Filtering Thresholds

Step	**Rule**	**Condition**
Quality rejection	R1	brightness < 70.0
Non-street rejection	R2	rail_share > 0.05 **AND** road_share < 0.02
	R3	water_share > 0.30 **OR** sky_share > 0.80 **AND** road_share < 0.02
	R4	road_share < 0.02 **AND** sky_share < 0.03
Pedestrian-accessible street	R5	sidewalk_bottom_anywhere > 0.10
	R6	sidewalk_lower_left > 0.10
	R7	sidewalk_lower_right > 0.10
	R8	pedestrian_area_share > 0.10
	R9	bike_lane_share > 0.10
	R10	sidewalk_left > 0.08 **OR** sidewalk_right > 0.08 **AND** sidewalk_bottom_anywhere > 0.05
Final label		Pedestrian-accesible street if **any of R5–R10 holds**; otherwise `other`

Table B.2: Rule-based image analysis and filtering thresholds

B.4. Step 4: Construction of Visual Categories

Algorithm 4 Construction of Visual Categories

```
Input: Filtered image dataset for all cities
for each city do
    Load clustered image dataset
    Keep only images labelled as pedestrian_roads
    Compute image-level variables: vehicle count, greenery presence, people count
end for
Combine all city datasets into one pooled dataset
Estimate global thresholds for each visual variable
Assign categorical labels to each image using predefined binning rules
Construct combined visual category profile from all assigned labels
Save combined dataset and city-specific outputs
Output: Binned dataset of pedestrian-accessible street images
```

Binning Rules for Visual Categories

Variable	**Method**	**Classes**	**Description**
Vehicle count	zero vs nonzero	2	Images are classified into `vehicle_none` and `vehicle_present` depending on whether the detected vehicle count equals zero or is greater than zero.
Greenery presence	quantile	3	Images are classified into `greenery_low`, `greenery_medium`, and `greenery_high` using quantile-based thresholds that divide the data into three equal groups.
People count	zero vs nonzero	2	Images are classified into `pedestrian_none` and `pedestrian_present` depending on whether the detected people count equals zero or is greater than zero.

Table B.3: Binning methods used to construct visual categories

C

Street-Level Images

C.1. Filtered Images

Figure C.1: Examples of images labelled using filtering thresholds

C.2. Visually Categorised Images

Figure C.2: Examples of images classified by greenery presence levels

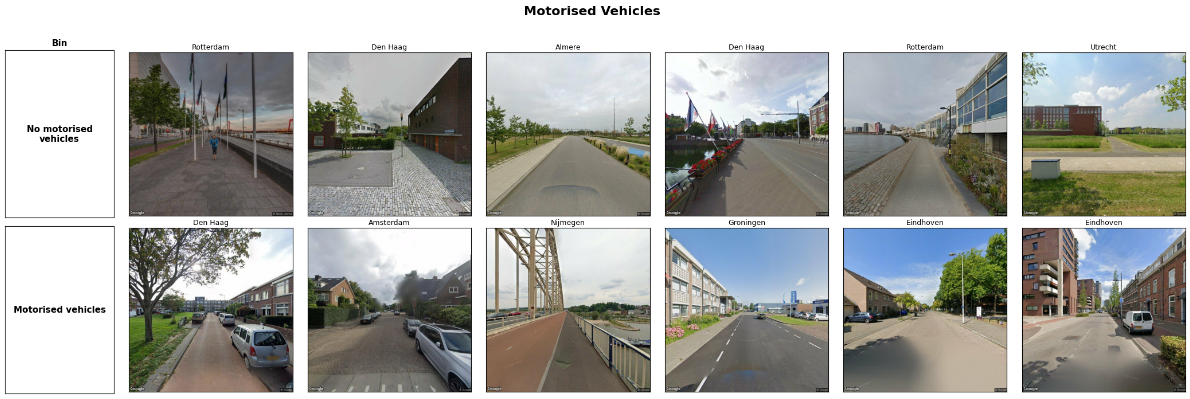


Figure C.3: Examples of images classified by motorised vehicle presence

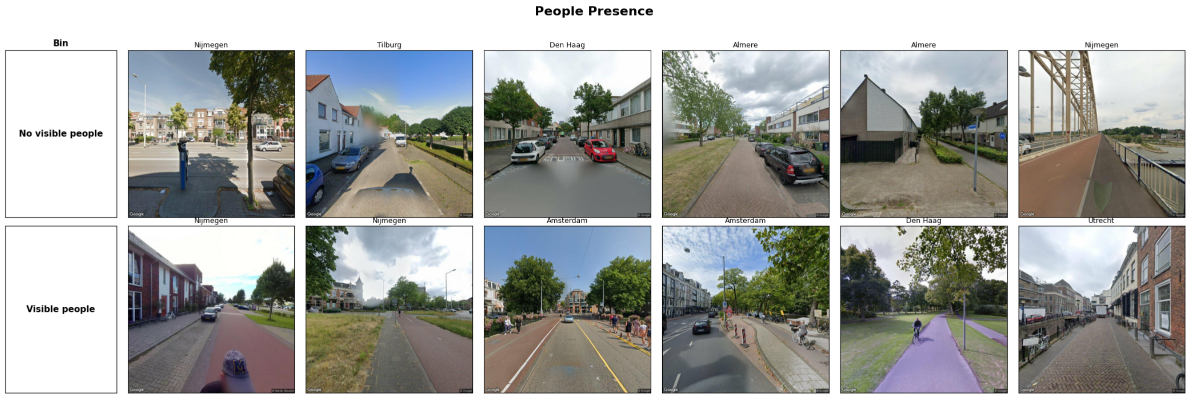


Figure C.4: Examples of images classified by people presence

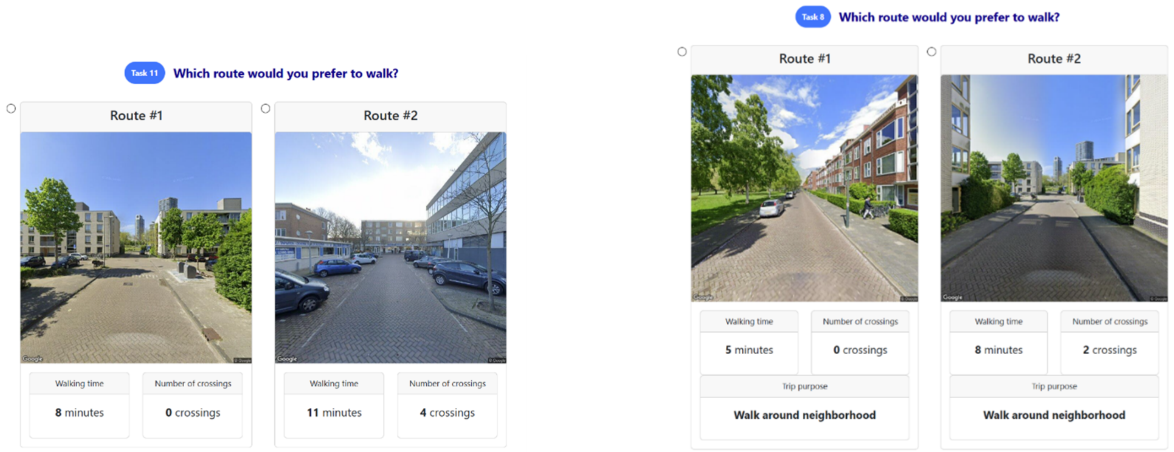
Figure C.5: Combined classification of images by vehicle presence, greenery presence, and people presence (first 6 out of 12)

Figure C.6: Combined classification of images by vehicle presence, greenery presence, and people presence (last 6 out of 12)

D

Survey Pilot Design

D.1. Choice Task Mock Designs



Task 11 Which route would you prefer to walk?

○ Route #1
Walking time: 8 minutes
Number of crossings: 0 crossings

○ Route #2
Walking time: 11 minutes
Number of crossings: 4 crossings

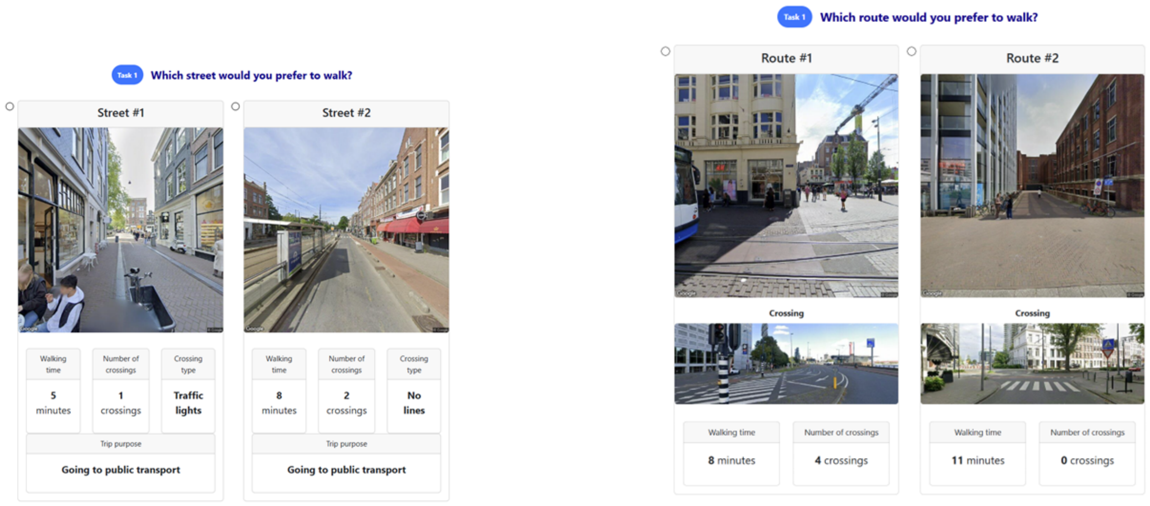
Task 8 Which route would you prefer to walk?

○ Route #1
Walking time: 5 minutes
Number of crossings: 0 crossings
Trip purpose: Walk around neighborhood

○ Route #2
Walking time: 8 minutes
Number of crossings: 2 crossings
Trip purpose: Walk around neighborhood

(a) Basic design: travel time and number of crossings

(b) Travel time and crossings with trip purpose



Task 1 Which street would you prefer to walk?

○ Street #1
Walking time: 5 minutes
Number of crossings: 1 crossings
Crossing type: Traffic lights
Trip purpose: Going to public transport

○ Street #2
Walking time: 8 minutes
Number of crossings: 2 crossings
Crossing type: No lines
Trip purpose: Going to public transport

Task 1 Which route would you prefer to walk?

○ Route #1
Crossing
Walking time: 8 minutes
Number of crossings: 4 crossings

○ Route #2
Crossing
Walking time: 11 minutes
Number of crossings: 0 crossings

(c) Crossing type represented as text

(d) Crossing type represented as images

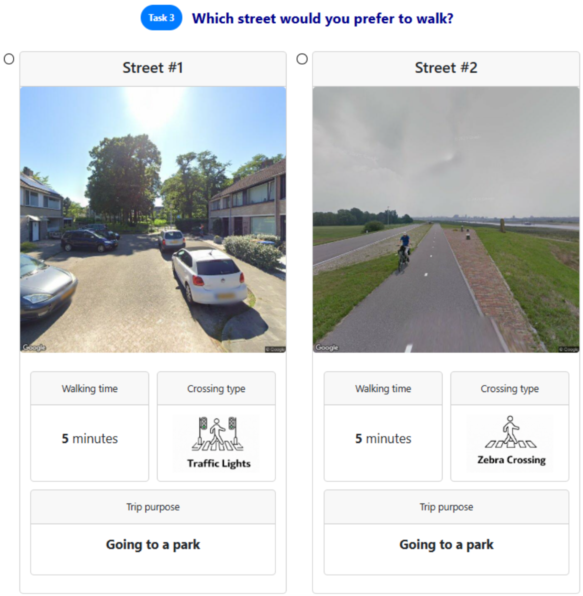
Figure D.1: Overview of mock choice task designs used during the initial development phase

D.2. Pilot Survey

The survey is developed in three stages: first creating mock designs, then testing a pilot survey, and lastly the final survey. The mock survey designs are created to explore how attributes can be presented and interpreted within the choice tasks. These designs include variations in layout, how attributes are shown and in which combinations they are included. The tested designs are shown in Appendix D.1. To evaluate these designs, talk-through sessions have been conducted with peers and colleagues in which participants complete the survey while explaining the reasoning behind their choices. This makes it possible to observe how respondents interpret the information, what they focus on, and how they arrive at their choices. From these sessions, the following points are identified as most relevant for the design of the choice tasks:

- **Number of crossings:** often ignored or not clearly understood by respondents.
- **Crossing type as text:** unclear what each type means, requiring respondents to refer back to instructions.
- **Crossing type as image:** creates conflicting information with the street image, leading to confusion about which image to focus on.
- **Trip purpose:** useful to include explicitly, as it helps to understand the context and recognise when it changes.

The pilot survey is developed based on the mock designs described in Section 3.4. The final pilot choice task design is shown in Figure D.2. In this design, respondents choose between two alternatives that differ in travel time, crossing type, and visual street characteristics. Crossing types are represented using icons to reduce potential conflicts with the street-level images.


Task 3 Which street would you prefer to walk?

○ Street #1

Walking time: 5 minutes

Crossing type: Traffic Lights

Trip purpose: Going to a park

○ Street #2

Walking time: 5 minutes

Crossing type: Zebra Crossing

Trip purpose: Going to a park

Figure D.2: Pilot choice task design using crossing icons and trip-purpose context

Pilot Choice Task Efficient Design

The experimental design for the pilot survey is generated using Ngene. The design consists of two alternatives per choice task, 24 rows in total, and is divided into 2 blocks. A D-efficient design is constructed under a multinomial logit specification using prior parameter values.

```
design
;alts = alt1, alt2
;rows = 24
;block = 2
;eff = (mnl,d)

;cond:
```

```
if(alt1.time = alt2.time and alt1.cross = alt2.cross, alt1.imgcat <> alt2.imgcat)

;model:
U(alt1) = b_time[-0.10] * time[5,8,11]
        + b_cross.dummy[-0.10|-0.20|-0.30] * cross[1,2,3,4]
        + b_img.dummy[-0.30|-0.25|-0.20|-0.15|-0.15|-0.10|-0.05] * imgcat[1,2,3,4,5,6,7,8]
/
U(alt2) = b_time * time
        + b_cross * cross
        + b_img * imgcat
$
```

Image Categories

The design includes 8 image categories, based on combinations of vehicle presence, greenery, and pedestrian presence. The ordering of these categories in the design corresponds to the following structure:

Category	Description	Utility
1	Vehicle high, greenery low, pedestrian high	-0.30
2	Vehicle high, greenery low, pedestrian low	-0.25
3	Vehicle high, greenery high, pedestrian high	-0.20
4	Vehicle high, greenery high, pedestrian low	-0.15
5	Vehicle low, greenery low, pedestrian high	-0.15
6	Vehicle low, greenery low, pedestrian low	-0.10
7	Vehicle low, greenery high, pedestrian high	-0.05
8	Vehicle low, greenery high, pedestrian low	0.00

Table D.1: Image category definitions and prior utilities

The most preferred category has utility 0 and corresponds to low vehicle presence, high greenery, and low pedestrian presence. The relative penalties reflect that high vehicle presence is strongly undesirable, low greenery is moderately undesirable, and high pedestrian presence is slightly undesirable.

Crossing Attribute

The crossing attribute consists of four levels:

Level	Crossing type	Utility
1	Zebra crossing	-0.10
2	Traffic light	-0.20
3	Unmarked crossing	-0.30
4	No crossing	0.00

Table D.2: Crossing types and prior utilities

The most preferred situation is having no crossing at all. Other crossing types are associated with increasing waiting times and thus negative utility.

E

Final Survey Design

The final survey consists of four main components: (1) an initial set of screening and demographic questions, (2) a set of stated choice tasks under different trip-purpose contexts, (3) perception questions based on individual street images, and (4) a final set of background and walking behaviour questions.

E.1. Survey Questions

Initial Questions

The initial questions are used to collect basic demographic information.

Question	Answers
What is your age group?	18–29 30–49 50–69 70 or older
What is your gender?	Male Female Other Prefer not to say
Do you currently live in the Netherlands?	Yes No
Are you able to walk independently for short distances?	Yes No

Table E.1: Initial demographic questions

Choice Tasks

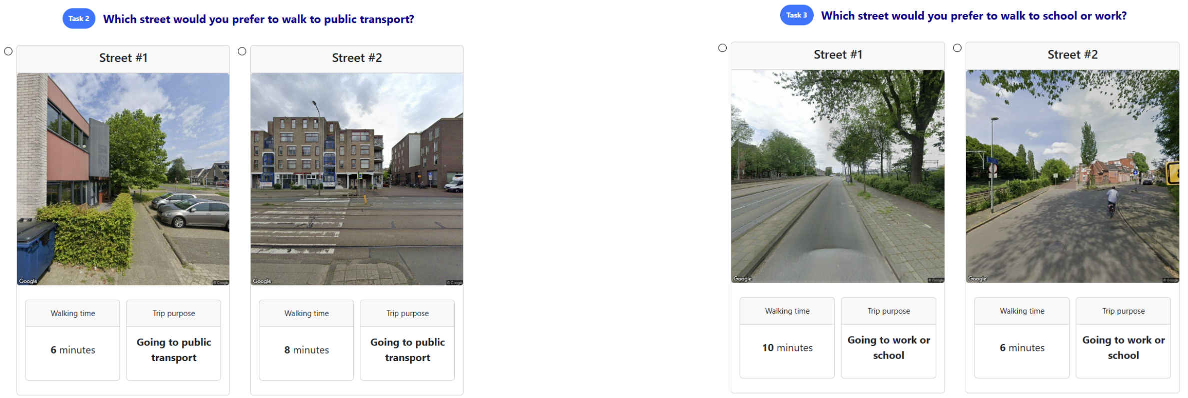
The choice task component consists of 12 stated choice questions. In each task, each alternative differs in terms of walking time and the visual appearance of the street. Each respondent completes three choice tasks for each of four trip-purpose contexts:

- Walking to public transport
- Walking to work or school
- Walking home
- Walking in free time

The respondents are given the following information for each choice task:
Imagine you are walking to {Trip Purpose} and you must choose between two directions to continue your walk.

Please assume:

- You are walking alone
- You are able to walk at your normal pace
- It is daytime and dry weather
- Walking time is the total time to reach your destination



Task 2 Which street would you prefer to walk to public transport?

○ Street #1
Walking time: 6 minutes
Trip purpose: Going to public transport

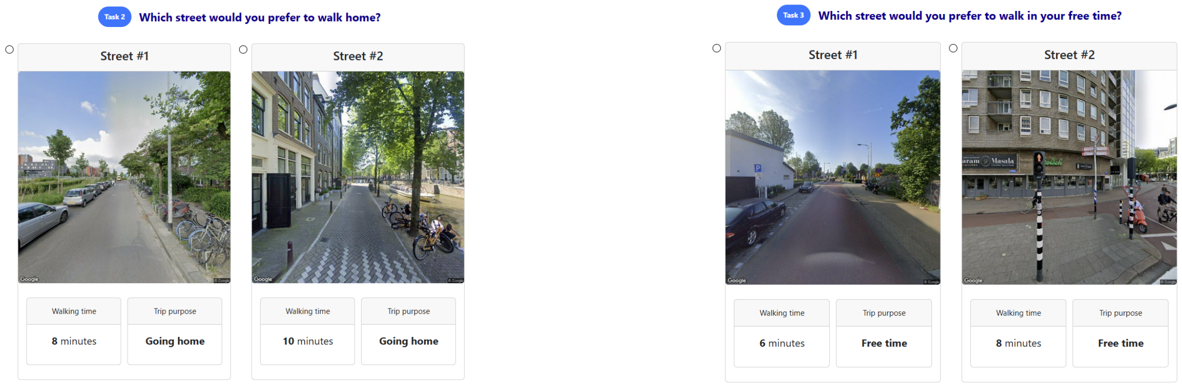
○ Street #2
Walking time: 8 minutes
Trip purpose: Going to public transport

Task 3 Which street would you prefer to walk to school or work?

○ Street #1
Walking time: 10 minutes
Trip purpose: Going to work or school

○ Street #2
Walking time: 6 minutes
Trip purpose: Going to work or school

(a) Walking to public transport **(b)** Walking to work or school



Task 2 Which street would you prefer to walk home?

○ Street #1
Walking time: 8 minutes
Trip purpose: Going home

○ Street #2
Walking time: 10 minutes
Trip purpose: Going home

Task 3 Which street would you prefer to walk in your free time?

○ Street #1
Walking time: 6 minutes
Trip purpose: Free time

○ Street #2
Walking time: 8 minutes
Trip purpose: Free time

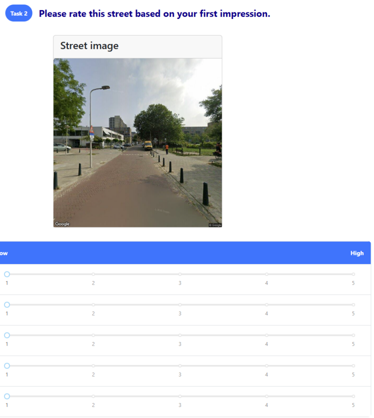
(c) Walking home **(d)** Walking in free time

Figure E.1: Overview of choice task designs used in the final survey

Perception Questions

In addition to the choice tasks, respondents are asked to evaluate four individual street images. These perception questions focus on the respondent's first impression of each image. Each image is rated on a five-point Likert scale for the following characteristics:

- Attractiveness: how pleasant and appealing the street feels to walk in
- Safety: how safe the respondent would feel walking in the street
- Greenery: the perceived amount of vegetation such as trees and plants
- Busyness (people): the perceived number of pedestrians
- Busyness (traffic): the perceived level of traffic



Task 2 Please rate this street based on your first impression.

Street image

?	Low				High
Attractiveness	○ 1	2	3	4	5
Safety	○ 1	2	3	4	5
Greenery	○ 1	2	3	4	5
Busyness (people)	○ 1	2	3	4	5
Busyness (traffic)	○ 1	2	3	4	5

Figure E.2: Example perception question used in the final survey

Final Questions

The final questions collect information on walking behaviour, preferences, and personal background characteristics.

Question	Answers
How often do you walk for trips?	Almost every day Several times per week About once per week Less than once per week Rarely or never
What is the main purpose you walk to places?	Going to or from public transport Going to work or school Leisure or free time Other
How long are your typical walking trips?	Less than 5 minutes 5–10 minutes 10–20 minutes 20–30 minutes More than 30 minutes
How do you feel about walking?	Strongly dislike Dislike Neutral Like Strongly like
How often do you feel unsafe while walking during the day?	Never Rarely Sometimes Often Always
What best describes the area you live in?	Large city Small city or town Village or rural area
Do you have access to a car in your household?	Yes No
What is your monthly household income before taxes?	Less than €1,500 €1,500–€2,500 €2,500–€3,500 €3,500–€5,000 More than €5,000 Prefer not to say
Which of the following best describes your place of residency (for most of your life)?	The Netherlands Another European country A country outside Europe Prefer not to say

Table E.2: Final background and walking behaviour questions

E.2. Final Choice Task Efficient Design

Prior Estimations

After 16 responses of the pilot survey, new priors were estimated. These priors where then scaled back to make sure the final priors are not dominated by the limited responses.

Parameter	Pilot estimate	Final prior (scaled)
Travel time	-0.278	-0.15
Vehicle present	-0.450	-0.25
Greenery low	-0.543	-0.30
Pedestrian present	-0.501	-0.25

Table E.3: Pilot estimates and adjusted priors for the final design

Image Category Priors

The design includes 12 image categories based on combinations of vehicle presence, greenery level, and pedestrian presence. Based on the priors in Table E.3, new priors for each image category are calculated. The most preferred category has a prior value of 0 and corresponds to no vehicle presence, high greenery, and no pedestrian presence. The remaining categories reflect decreasing preference due to the presence of vehicles, reduced greenery, and increased pedestrian density.

Category	**Description**	**Prior**
1	Vehicle present, greenery low, pedestrian present	-0.80
2	Vehicle present, greenery low, pedestrian none	-0.55
3	Vehicle present, greenery medium, pedestrian present	-0.65
4	Vehicle present, greenery medium, pedestrian none	-0.40
5	Vehicle present, greenery high, pedestrian present	-0.50
6	Vehicle present, greenery high, pedestrian none	-0.25
7	Vehicle none, greenery low, pedestrian present	-0.55
8	Vehicle none, greenery low, pedestrian none	-0.30
9	Vehicle none, greenery medium, pedestrian present	-0.40
10	Vehicle none, greenery medium, pedestrian none	-0.15
11	Vehicle none, greenery high, pedestrian present	-0.25
12	Vehicle none, greenery high, pedestrian none	0.00

Table E.4: Image category definitions and priors for the image categories of the final survey

E.2.1. Experimental Design

The efficient experimental design for the final survey is generated using Ngene. The design consists of two alternatives per choice task, 36 rows in total, and is divided into 3 blocks. A D-efficient design is constructed under a multinomial logit specification using prior parameter values derived from the pilot survey.

```
design
;alts = alt1, alt2
;rows = 36
;block = 3
;eff = (mnl,d)

;cond:
if(alt1.time = alt2.time, alt1.imgcat <> alt2.imgcat)

;model:
U(alt1) = b_time[-0.15] * time[6,8,10]
        + b_img.dummy[-0.80|-0.55|-0.65|-0.40|-0.50|-0.25|-0.55|-0.30|-0.40|-0.15|-0.25] *
            imgcat[1,2,3,4,5,6,7,8,9,10,11,12]
/
U(alt2) = b_time * time
        + b_img * imgcat
```

```
$
```

Table E.5: Final experimental design (Ngene output)

Design	Choice situation	alt1.time	alt1.imgcat	alt2.time	alt2.imgcat	Block
1	1	8	1	10	12	3
1	2	8	9	6	12	3
1	3	10	9	6	12	2
1	4	6	5	6	6	3
1	5	10	1	10	3	2
1	6	8	9	8	3	3
1	7	10	8	8	11	1
1	8	8	7	6	4	2
1	9	6	10	10	5	3
1	10	6	2	8	7	3
1	11	10	6	6	2	3
1	12	10	12	6	8	2
1	13	10	10	6	8	1
1	14	8	7	8	5	1
1	15	8	2	6	4	2
1	16	10	4	10	11	1
1	17	8	8	8	5	1
1	18	8	2	8	4	3
1	19	6	5	6	1	2
1	20	8	4	10	10	2
1	21	10	8	8	2	2
1	22	8	12	8	4	1
1	23	10	4	10	9	1
1	24	8	11	10	4	3
1	25	8	3	10	10	3
1	26	10	11	8	10	3
1	27	6	6	6	7	1
1	28	6	3	10	6	1
1	29	6	1	10	9	1
1	30	6	5	10	12	2
1	31	10	3	6	7	2
1	32	6	10	6	1	1
1	33	10	12	8	6	2
1	34	10	11	6	9	3
1	35	8	7	6	8	1
1	36	10	6	10	1	2

E.3. Pseudocode for Choice Task Design Construction

This section describes the main steps used to translate the Ngene design into the final respondent-specific choice tasks. The procedure assigns attribute combinations to trip-purpose contexts and links these to street-level images.

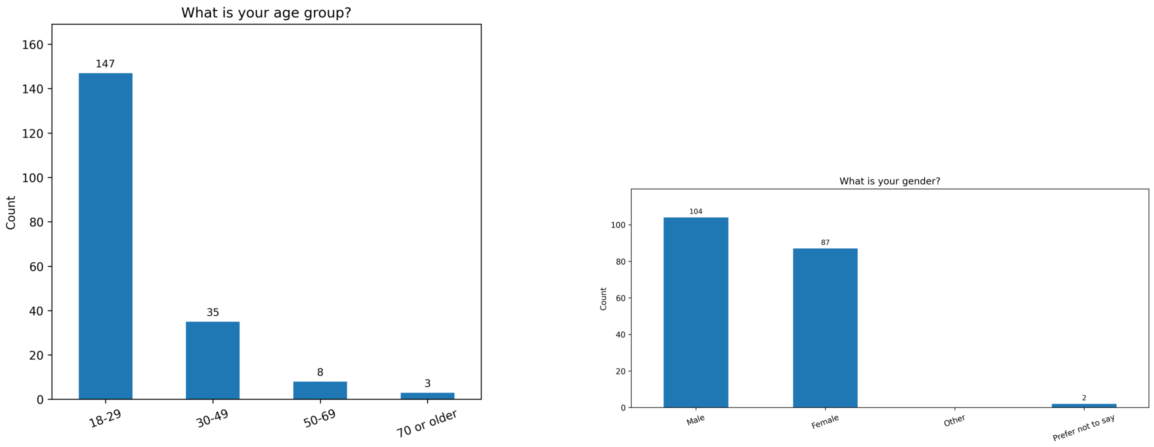
Algorithm 5 Construction of Respondent-Specific Choice Tasks

```
Input: Ngene design, image metadata, number of respondents
Read the Ngene design table
Keep the attributes for each alternative:
    walking time, image category, and block number
Validate the design:
    allowed time levels
    allowed image categories
    12 tasks per block
Read the image metadata
Map each image to one of the predefined image categories
Group images into pools by category
for each respondent do
    Assign one design block to the respondent (cycling over blocks)
    Select the 12 tasks of that block
    Order the tasks from 1 to 12 within the block
    Divide the 12 tasks into 4 groups:
        group 1: tasks 1–3
        group 2: tasks 4–6
        group 3: tasks 7–9
        group 4: tasks 10–12
    Assign the 4 groups to the 4 trip-purpose contexts:
        context 1: public transport
        context 2: work or school
        context 3: home
        context 4: free time
    Apply cyclic rotation across respondents:
        respondent 1: groups → contexts (1, 2, 3, 4)
        respondent 2: groups → contexts (2, 3, 4, 1)
        respondent 3: groups → contexts (3, 4, 1, 2)
        respondent 4: groups → contexts (4, 1, 2, 3)
        repeat this cycle for all respondents
    for each task do
        Identify the group to which the task belongs
        Assign the corresponding (rotated) trip-purpose context
        Read the image category of alternative 1 and alternative 2
        Select one image from the correct pool for each alternative
        Avoid using the same image twice within one task
        Avoid reusing the same image within one respondent where possible
        Combine:
            selected images
            walking time levels
            trip-purpose label
        Randomly flip the left-right order of the two alternatives
    end for
end for
Separate the final tasks by trip-purpose context
Save the resulting choice tasks as survey design files
Output: Final respondent-specific choice tasks
```

F

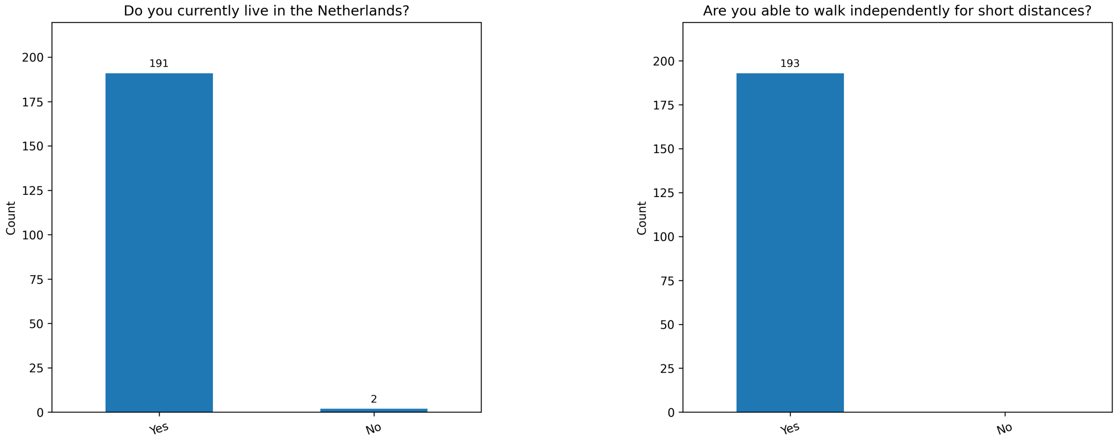
Socio-Demographics

This appendix provides an overview of the socio-demographic characteristics and walking-related background variables of the survey respondents. These plots are included to give additional context to the respondent sample used in the stated preference experiment.



(a) Age group

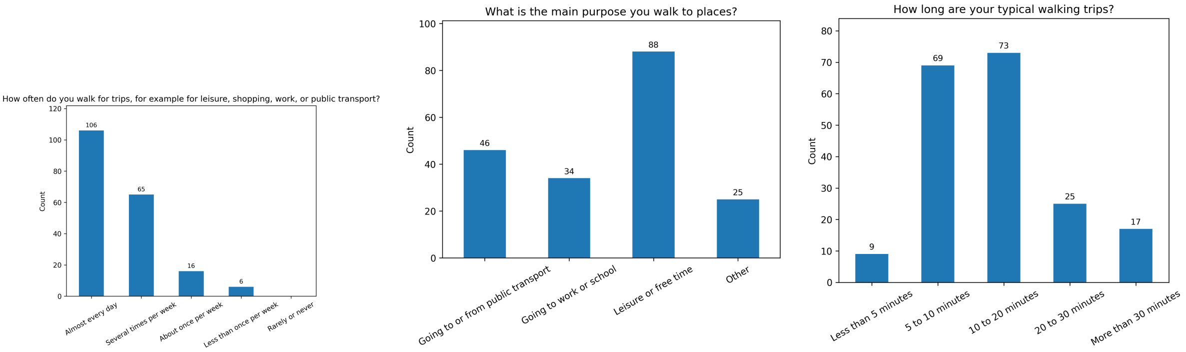
(b) Gender



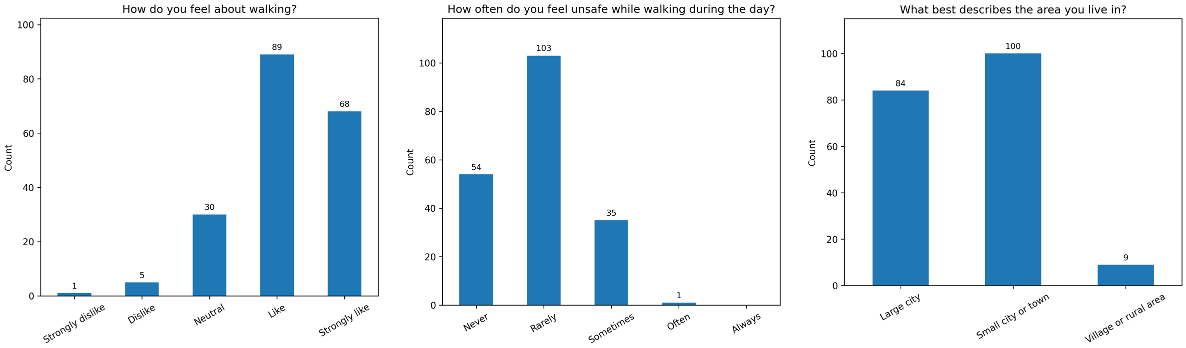
(c) Currently living in the Netherlands

(d) Able to walk independently

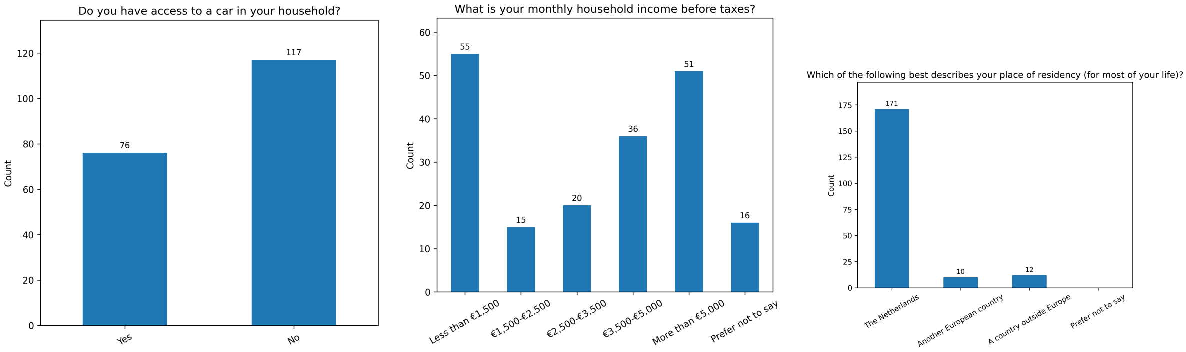
Figure F.1: Distribution of responses for the screening and demographic questions.



(a) Walking frequency **(b)** Main walking purpose **(c)** Typical walking duration



(d) Attitude towards walking **(e)** Feeling unsafe during the day **(f)** Living environment



(g) Car access **(h)** Household income **(i)** Place of residency

Figure F.2: Distribution of responses for the background questions.

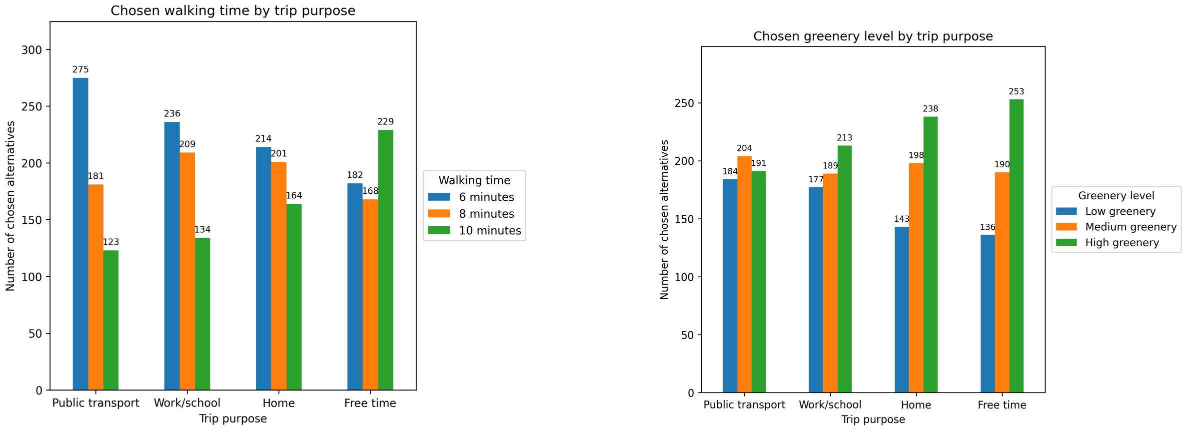
G

Additional Results

This appendix contains additional results that support the main results chapter. These results are not discussed in detail in the main text, but are included to make the analysis more transparent.

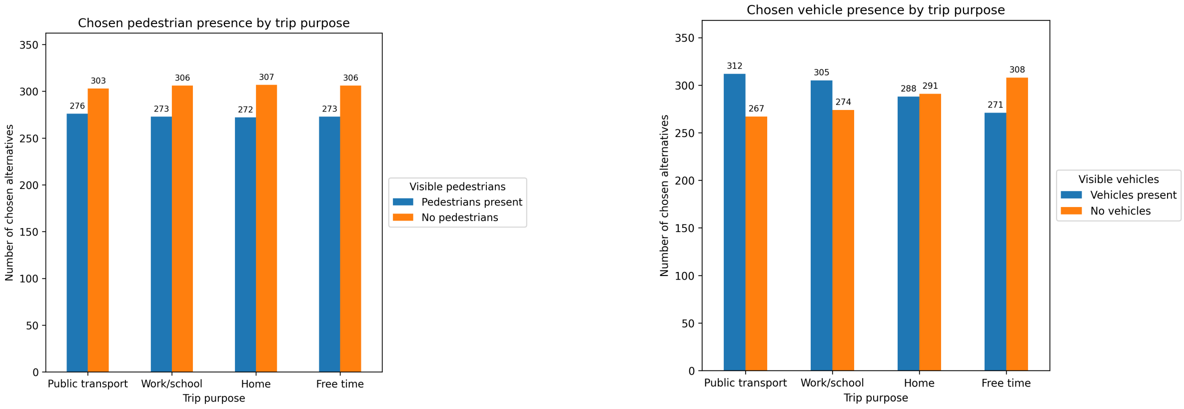
G.1. Additional Choice Task Patterns

The main results chapter shows the distribution of chosen walking times and greenery levels. Figure G.1 shows the additional choice task patterns for pedestrian presence and vehicle presence.



(a) Chosen walking time by trip purpose

(b) Chosen greenery level by trip purpose



(c) Chosen pedestrian presence by trip purpose

(d) Chosen vehicle presence by trip purpose

Figure G.1: Distribution of chosen route attributes by trip purpose.

G.2. Alternative Pixel-Share Model Specifications

This appendix reports the alternative pixel-share MNL specifications that were tested as robustness checks. These include reduced specifications and specifications with an alternative-specific constant. The main results chapter uses the full pixel-share MNL specification to maintain comparability across trip purposes and model types.

	Best configuration				
Walking Purpose	**Combined**	**Public Transport**	**Work or School**	**Home**	**Free Time**
Model information					
Model type	RUM-MNL	RUM-MNL	RUM-MNL	RUM-MNL	RUM-MNL
No. parameters	4	4	3	5	3
Train set					
N	1853	463	469	464	457
Log-likelihood	-1177.44	-262.18	-279.85	-290.87	-294.19
ρ^2	0.083	0.183	0.139	0.096	0.071
Cross-entropy	0.635	0.566	0.597	0.627	0.644
Hit-rate	0.648	0.708	0.682	0.659	0.637
BIC	2384.97	548.91	578.14	612.45	606.76
Test set					
N	463	116	110	115	122
Log-likelihood	-276.54	-65.95	-60.93	-63.70	-76.99
ρ^2	0.138	0.180	0.201	0.201	0.090
Cross-entropy	0.597	0.569	0.554	0.554	0.631
Hit-rate	0.687	0.698	0.718	0.730	0.623
BIC	577.64	150.91	135.96	151.12	168.40
Parameter estimates					
β_{time}	-0.276***	-0.492***	-0.406***	-0.301***	0.021
β_{sidewalk}	-0.010	-0.017	–	–	–
$\beta_{\text{pedestrian area}}$	–	–	–	0.010	0.008
β_{vehicles}	–	0.001	–	-0.010	–
$\beta_{\text{vegetation}}$	0.025***	0.019***	0.023***	0.029***	0.031***
β_{people}	0.057*	–	0.091	0.049	–

Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table G.1: Performance and parameter comparison for the pixel-share MNL models.

Walking Purpose	Configuration with alternative-specific constant Com-bined	 Public Transport	 Work or School	 Home	 Free Time
Model information					
Model type	RUM-MNL	RUM-MNL	RUM-MNL	RUM-MNL	RUM-MNL
No. parameters	7	7	7	7	7
Train set					
N	1853	463	469	464	457
Log-likelihood	-1169.85	-259.42	-273.51	-288.50	-293.71
ρ^2	0.089	0.192	0.159	0.103	0.073
Cross-entropy	0.631	0.560	0.583	0.622	0.643
Hit-rate	0.644	0.713	0.695	0.657	0.621
BIC	2392.37	561.80	590.07	619.99	630.30
Test set					
N	463	116	110	115	122
Log-likelihood	-276.96	-65.94	-64.79	-64.51	-77.45
ρ^2	0.137	0.180	0.150	0.191	0.084
Cross-entropy	0.598	0.568	0.589	0.561	0.635
Hit-rate	0.683	0.690	0.709	0.722	0.631
BIC	596.88	165.16	162.48	162.24	188.52
Parameter estimates					
$\mathrm{ASC}_{\mathrm{ALT1}}$	-0.011	0.056	-0.055	-0.204*	0.086
β_{time}	-0.286***	-0.512***	-0.444***	-0.307***	0.023
$\beta_{\mathrm{sidewalk}}$	0.001	-0.002	-0.011	0.012	0.000
$\beta_{\mathrm{pedestrian\ area}}$	0.011***	0.014*	0.015*	0.012	0.008
$\beta_{\mathrm{vehicles}}$	-0.001	0.016	-0.017	-0.010	0.006
$\beta_{\mathrm{vegetation}}$	0.026***	0.021***	0.026***	0.030***	0.031***
β_{people}	0.037	0.029	0.059	0.047	-0.006

Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Note: $\mathrm{ASC}_{\mathrm{ALT1}}$ is the alternative-specific constant for alternative 1. The constant for the reference alternative is fixed to zero.

Table G.2: Performance and parameter estimates for the pixel-share MNL models with an alternative-specific constant.

G.3. CV-DCM Hyperparameter Tuning Results

This appendix summarises the hyperparameter tuning procedure for the CV-DCM. Rather than reporting all individual optimisation runs, it reports the search configurations that were tested and the final selected settings. The selected settings were based on the lowest validation cross-entropy, while also considering stability across seeds where relevant.

Search step	Model scope	Optimizer(s)	Learning rate(s)	Batch size(s)	Weight decay
Optimizer comparison	Combined	SGD, Adam	$1\mathrm{e}{-6}$, $1\mathrm{e}{-5}$	5, 20, 24	0.1
SGD coarse search	Combined	SGD	$1\mathrm{e}{-5}$, $2\mathrm{e}{-5}$	5, 10, 15	0.0, 0.1, 0.2
SGD fine search	Combined	SGD	$8\mathrm{e}{-6}$, $1\mathrm{e}{-5}$, $1.2\mathrm{e}{-5}$	8, 10, 12	0.01, 0.05, 0.1
Trip-purpose search	Trip-specific	SGD	$1\mathrm{e}{-5}$, $2\mathrm{e}{-5}$, $3\mathrm{e}{-5}$	5, 8, 10	0.05, 0.1, 0.2, 0.3

Table G.3: Overview of the CV-DCM hyperparameter search procedure.

Model	Optimizer	LR	Batch	WD	CE
Combined	SGD	$1\mathrm{e}{-5}$	10	0.05	0.5579
Free time	SGD	$2\mathrm{e}{-5}$	8	0.3	0.5865
Going home	SGD	$1\mathrm{e}{-5}$	8	0.1	0.4689
Going to public transport	SGD	$2\mathrm{e}{-5}$	5	0.2	0.5495
Going to work or school	SGD	$2\mathrm{e}{-5}$	8	0.2	0.5703

Table G.4: Selected CV-DCM hyperparameter configurations based on validation cross-entropy.

	Trip-purpose-specific CV-DCM configuration			
Metric	**Free**	**Home**	**PT**	**Work**
Model information				
Model type	CV-DCM	CV-DCM	CV-DCM	CV-DCM
No. parameters	86m	86m	86m	86m
Optimizer	SGD	SGD	SGD	SGD
Learning rate	$2\mathrm{e}{-5}$	$1\mathrm{e}{-5}$	$2\mathrm{e}{-5}$	$2\mathrm{e}{-5}$
Batch size	8	8	5	8
L2	0.3	0.1	0.2	0.2
Train set				
N	457	464	463	469
Log-likelihood	-202.11	-123.02	-200.47	-212.43
ρ^2	0.359	0.618	0.378	0.345
Cross-entropy	0.444	0.265	0.431	0.454
Hit-rate	0.807	0.927	0.826	0.799
BIC	–	–	–	–
Test set				
N	122	115	116	110
Log-likelihood	-72.73	-53.45	-62.64	-63.30
ρ^2	0.154	0.324	0.207	0.177
Cross-entropy	0.587	0.469	0.550	0.570
Hit-rate	0.702	0.781	0.728	0.712
BIC	–	–	–	–
Parameter estimates				
β_{time}	-0.092	-0.262	-0.409	-0.311
Image utility	-	-	-	-

Table G.5: Performance and hyperparameter configuration for the trip-purpose-specific CV-DCM models.

G.4. Same-Street Qualitative Validation Results

This appendix presents additional same-street comparisons used for the qualitative validation of the CV-DCM. These comparisons show how predicted image utilities vary between images taken along the same street. They are included as supporting material for the interpretation of the image utilities, while the main results chapter focuses on the overall validation findings.

Figure G.2: Same-street comparison of predicted image utilities for all trip purposes.

G.5. AI Image Augmentations

This appendix reports the prompts used to generate the AI-edited images and the images with utilities. The edits were designed to change one visual element at a time while keeping the original street, camera angle, lighting, buildings, road layout, and overall composition unchanged.

The following prompts were used:

- **More greenery:** Edit the input street-level image realistically. Keep the same street, same camera angle, same lighting, same buildings, same road layout, and same overall composition. Only change the scene by adding more visible greenery, such as trees, shrubs, grass, or planted areas. Do not change anything else.
- **More traffic:** Edit the input street-level image realistically. Keep the same street, same camera angle, same lighting, same buildings, same road layout, and same overall composition. Only change the scene by adding more cars and visible traffic. Do not change anything else.
- **Trash and litter:** Edit the input street-level image realistically. Keep the same street, same camera angle, same lighting, same buildings, same road layout, and same overall composition. Only change the scene by adding multiple clearly visible trash bags and litter on the ground. Do not change anything else.
- **Pedestrian amenities:** Edit the input street-level image realistically. Keep the same street, same camera angle, same lighting, same buildings, same road layout, and same overall composition. Only change the scene by adding more pedestrian amenities, especially benches. Do not change anything else.

Figure G.3 shows the full set of AI-edited image scenarios. The main results chapter only reports the distribution of utility changes across the edited images.

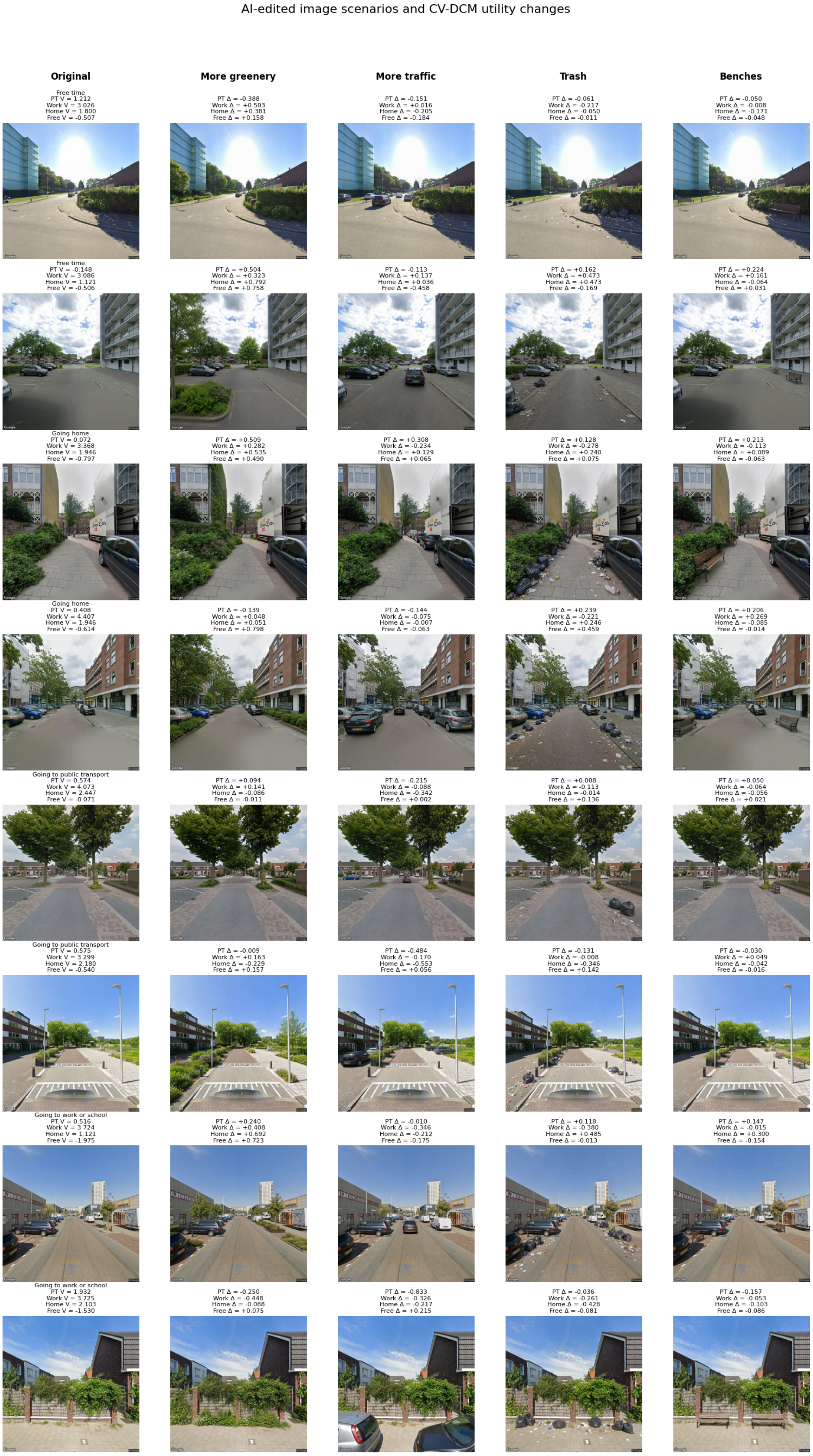
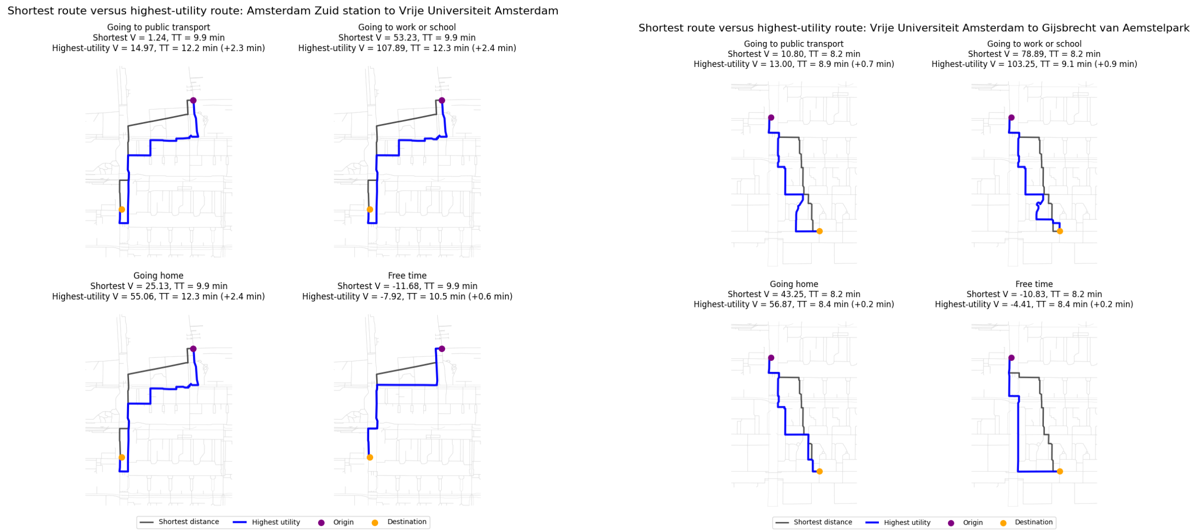


Figure G.3: AI-edited image scenarios and changes in predicted CV-DCM image utility.

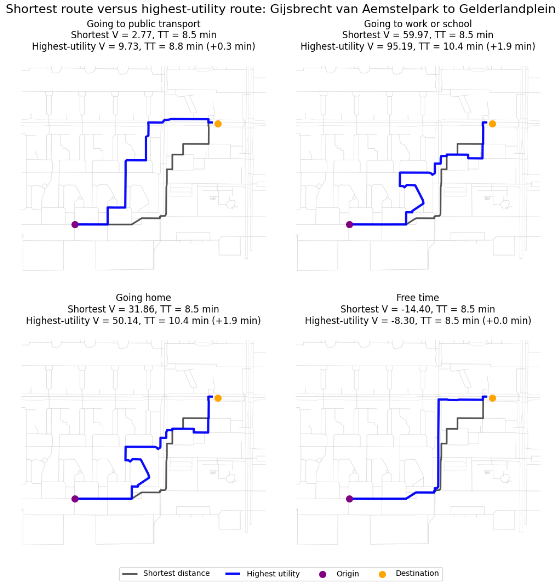
G.6. Spatial Application Route Comparisons

This appendix presents the route comparisons for the selected origin-destination pairs in the Amsterdam-Zuid spatial application. The selected routes connect Amsterdam Zuid station to Vrije Universiteit Amsterdam, Vrije Universiteit Amsterdam to the Gijsbrecht van Aemstelpark, and the Gijsbrecht van Aemstelpark to Gelderlandplein. For each origin-destination pair, the shortest-distance route is compared with the route that maximises total segment utility within the destination-directed route corridor.



(a) Amsterdam Zuid station to Vrije Universiteit Amsterdam.

(b) Vrije Universiteit Amsterdam to the Gijsbrecht van Aemstelpark.



(c) Gijsbrecht van Aemstelpark to Gelderlandplein.

Figure G.4: Comparison between the shortest-distance route and the highest-utility route for the selected origin-destination pairs in Amsterdam-Zuid.