Understanding cycling route choice behaviour through street-level images and computer vision-enriched discrete choice models

R.W. Terra [a,b], F. Garrido-Valenzuela [a,b], O. Cats[b,c], S. van Cranenburgh [a,b]

[a]*Transport and Logistics group, Faculty of Technology Policy and Management, Delft University of Technology, The Netherlands*
[b]*CityAI Lab, Delft University of Technology, The Netherlands*
[c]*Department of Transport & Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, The Netherlands*

Date issued: 12/12/2024

Abstract

This study investigates cyclists' preferences for cycling environments, such as lane type, pavement type, and lane width. We conducted a stated route choice experiment where respondents evaluated alternatives varying in cycling environment, travel time and the number of traffic lights. Unlike previous studies, we used thousands of street-level images to represent the cycling environment, providing rich information on safety and quality that text or numbers cannot convey. We analysed the data using recently proposed computer vision-enriched discrete choice models, which integrate computer vision into traditional discrete choice models. Thereby, we can infer cyclist trade-offs and estimate willingness-to-pay values. Results show that cycling environments strongly influence route choice, with cyclists preferring green areas and separated cycling lanes. For an 11-minute trip, cyclists are willing to take a 1.5-minute detour for a separate lane instead of a mixed-traffic road. These findings offer insights for designing cycling environments that align with cyclists' preferences.

Keywords: discrete choice models, computer vision, street-level images, stated choice experiment, cycling route behaviour, cycling environment

1 Introduction

Numerous transport policies aim to promote cycling due to its significant contributions to health, accessibility, sustainability, and liveability (Dutch Cycling Embassy, 2018). Furthermore, cycling is widely regarded as a key solution for achieving the 2015 Paris Climate Agreement targets, including a 40% reduction in CO_2 emissions by 2030 (European Union, 2023; I&W, 2020).

One way for governments to promote cycling is by providing cycling environments that align with cyclists' preferences. Fosgerau et al. (2023) find that a good cycling environment pulls travellers out of their cars for short-distance trips. Furthermore, various studies show that cyclists prefer cycling lanes, wider paths (Gössling & McRae, 2022), and separated lanes over shared ones (Kaplan & Prato, 2015). Additionally, road surface (Zimmermann et al., 2017), tram tracks (Kaplan & Prato, 2015), nearby parking spaces (Gössling & McRae, 2022), and built environment factors such as buildings, greenery, and public facilities are found to impact on route choice behaviour (Heinen et al., 2010; Wang et al., 2016; Yang et al., 2019).

Despite the above-mentioned body of knowledge, the understanding of cyclists' preferences for cycling environments remains incomplete. This gap partly arises from the reliance on stated choice experiments that describe cycling environments using text or numerical values. For example, in a stated choice experiment by Y. Liu et al. (2020), attributes such as "the average number of building floors along streets", "the width of the bicycle path", "the amount of street greenery", "the crowdedness", and "the density of street lamps" were described using text and numbers. Using text and numbers to describe a cycling environment may leave respondents struggling to fully grasp the characteristics of the described cycling environment. As a result, the preference parameters estimated

for the cycling environment from such an experiment may be biased.

Unlike tabular data, street-level images are particularly adept at encoding information about the cycling environment. Specifically, images have two important advantages over traditional textual descriptions or numbers. Firstly, the images can embed important information about the cycling environment, such as scale, safety, or quality, that is difficult to convey otherwise. Secondly, humans are adept at extracting and processing information from images. Moreover, recent advances in computer vision have opened up new ways of handling and extracting information from images (Pinker, 1988). Capitalising on these developments, a stream of literature has emerged that explores human perceptions of urban environments. For examples, see Dubey et al. (2016) and L. Liu et al. (2017), and refer to Zhang et al. (2024) for a comprehensive review. Within this stream of the literature, several studies also looked into cyclists' perceptions of cycling environments using street-level imagery (Costa, Azevedo, et al., 2024; Costa, Marques, et al., 2024; Ye et al., 2024). But perceptions and preferences are distinct concepts. Perceptions are subjective interpretations of sensory stimuli, which may influence but do not necessarily determine individuals' choices (Wade & Swanston, 2013). In contrast, preferences are grounded in the theory of choice behaviour (Samuelson, 1938; Luce, 1959; Lancaster,1966) and govern what people choose and how they make trade-offs. As yet, no cyclist route preference study has been conducted that uses street-level images to encode the cycling environment.

This study aims to fill this gap. It investigates cyclists' preferences for the cycling environment through a stated choice experiment in which the cycling environment is represented using street-level images. To do so, we have developed and administered a novel stated choice experiment in which respondents were presented with alternatives that varied in terms of the cycling environments (street-level images), travel time and the number of traffic lights. The survey incorporated images sourced from a cycling infrastructure database maintained by the municipality of Rotterdam, The Netherlands. We then analyse our stated choice data using recently proposed model computer vision-enriched discrete choice models (CV-DCMs) (van Cranenburgh & Garrido-Valenzuela, 2025). CV-DCMs integrate computer vision into traditional discrete choice models. Because they are grounded in Random Utility Maximisation (RUM) principles, they enable inferring cyclist trade-offs as well as willingness-to-pay estimates. Finally, we apply our estimated model to glean insights into the spatial distribution of the cycling environment, using Rotterdam as the case study.

The remainder of this paper is organised as follows. Section 2 introduces our data collection effort, including the design of the stated choice experiment. Section 3 presents the methodology and revisits the CV-DCMs. Section 4 presents the training/estimation results and the model application to our case study: the cycling environment in the municipality of Rotterdam. Finally, Section 5 provides a conclusion and discusses avenues for further research.

2 Data collection

This section outlines the process of data collection in two parts. First the cycling environment data collection in Section 2.1 details the collection and categorisation of street-level images and the associated cycling environment data. Section 2.2 outlines the generation of the stated choice experiment, detailing the attributes included, the underlying design, and respondent sampling. To provide a comprehensive overview, Figure 1 illustrates the workflow.

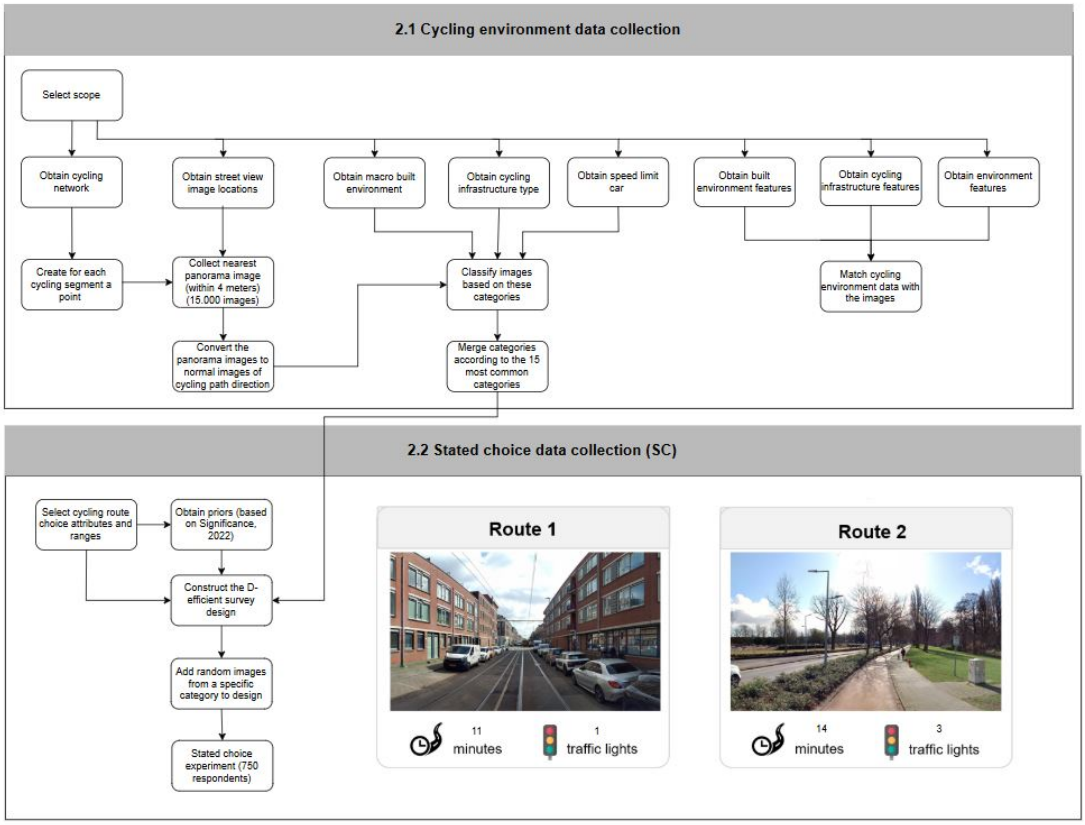


Figure 1: Methodology of the data collection

2.1 Cycling environment data collection

Our data collection process started with collecting street-level images of the cycling infrastructure. We obtained our images from the database of cycling infrastructure images of the municipality of Rotterdam (Cyclomedia, 2023). To extract a subset of relevant images, we merged their locations onto the cycling network of Fietsersbond (2020). Specifically, for each street segment, the nearest street-level image within a 4-metre radius was retrieved, resulting in approximately 13k panorama images for further processing. These panorama images were then converted to 90-degree forward-facing images aligned with the cycling direction.

Next, we analysed and categorised images according to cycling infrastructure, built environment, and vehicle speed. The decision was taken to categorise the images according to the 15 most common categories. This analysis provided the basis for selecting images for the stated choice experiment. Table 1 shows the attributes and their corresponding levels.

Table 1: Attributes and corresponding levels

Cycling Infrastructure	**Built Environment**	**Speed Limit**
Solitary cycling lane	Recreational area	Not applicable
Separated cycling lane	Residential area	30 km/h
Normal road	Neighbourhood access	50 km/h
Cycling suggestion lane	Main road	
	Industrial area	

In addition, we merged the image data with cycling environment data, including: built environment features, such as trees and buildings; Cycling infrastructure features, such as pavement type and lane width; and Traffic and environment features, such as car speed, weather and lightning. Street-level image captures times enabled matching with weather data of Visual Crossing (2024). Table 2 lists the attributes and their data sources. Merging these environmental features enabled us to design the stated choice experiment in such a way that meaningful trade-offs were observed. Note that for merging, different distance thresholds were used for different features.

For example, buildings and water were assigned if present within 20m of the image, while for attributes like grass, plants, parking, and trams, a 10m buffer was used. Trees were counted along the segment. The final image database contains approximately 13k images, with fifteen categories based on combinations of the Cycling Infrastructure, the Speed Limit, and the Built Environment.

Table 2: Collected attributes and their linking method

Attribute	Levels	Source	Linking method
Built Environment			
Trees	Count	BGT	Count within 20m buffer
Grass	Yes/No	BGT	Present within 10m
Plants	Yes/No	BGT	Present within 10m
Water	Yes/No	BGT	Present within 20m
Buildings	Yes/No	BGT	Present within 20m
Industrial buildings	Yes/No	BGT	Present within 20m
Cycling Infrastructure			
Infrastructure type	Categories	Fietsersbond	Initial network
Pavement type	Asphalt/Clinkers	Fietsersbond	Initial network
Parking	Yes/No	BGT	Present within 10m
Width	Metres	BGT	Present within 10m
Colour	Red/Not red	BGT	Present within 10m
Tram	Yes/No	BGT	Present within 10m
Traffic and Environment			
Car speed	30/50 km/h	Rijkswaterstaat	Present within 10m
Weather	Sunny/Cloudy	Visual Crossing	Matched with image timestamp
Lighting	Brightness score	Calculated	Computed from image

2.2 Stated choice data collection (SC)

2.2.1 Selected attributes and ranges

In our stated choice experiment, respondents were presented with a route choice. Based on the literature of Bernardi et al. (2018), Ton et al. (2017), and Verhoeven et al. (2018), expert input, and their relevance to cycling decisions and policy, we conceptualised the cycling route alternatives using three attributes: the cycling environment (represented using a street-level image), the travel time and the number of traffic lights. We used 8, 11, and 14 minutes for the travel time, representing realistic durations for urban cycling trips; for traffic lights, we used 1, 2, and 3.

2.2.2 Efficient design

An efficient design was chosen to maximise the information obtained from a stated choice experiment, resulting in more reliable parameter estimates for a given number of observations (Rose & Bliemer, 2009). Priors for the model parameters were based on findings from Significance (Significance, 2022), as shown in Table 3. The design was created using Ngene software (ChoiceMetrics, 2024), generating 30 choice tasks. In the experiment, images were randomly sampled from the image database based on their category. Furthermore, to manage cognitive load, each respondent was presented with 15 choice situations. The question order and the position of alternatives (left/right) were randomised.

Table 3: Prior beta parameters. Adapted from Significance (2022)

Parameter	Beta Value	Cycling Infrastructure	Speed Limit	Built Environment
b_time	-0.2707			
b_trafficlights	-0.1330			
b_img.dummy[0]	-1.002	Normal road	50	Access road
b_img.dummy[1]	-0.921	Normal road	50	Industrial
b_img.dummy[2]	-0.850	Normal road	50	Residential
b_img.dummy[3]	-0.720	Cycling suggestion lane	50	Access road
b_img.dummy[4]	-0.663	Normal road	30	Industrial
b_img.dummy[5]	-0.582	Normal road	30	Main road
b_img.dummy[6]	-0.512	Normal road	30	Access road
b_img.dummy[7]	-0.471	Cycling suggestion lane	30	Access road
b_img.dummy[8]	-0.466	Normal road	30	Residential
b_img.dummy[9]	-0.390	Separated cycling lane	N/A	Main road
b_img.dummy[10]	-0.319	Separated cycling lane	N/A	Access road
b_img.dummy[11]	-0.273	Separated cycling lane	N/A	Residential
b_img.dummy[12]	-0.116	Solitary cycling lane	N/A	Main road
b_img.dummy[13]	-0.046	Solitary cycling lane	N/A	Residential
b_img.dummy[14]	0	Solitary cycling lane	N/A	Recreational

2.2.3 Survey design

The survey was developed using Plotly (2024). The survey comprised three parts: 1) a Stated choice experiment; 2) Perception questions involving the rating images on traffic safety, social safety, and aesthetics; and 3) Questions about socio-demographic characteristics and cycling behaviour. The survey can be accessed at: http://cycling-route-survey.tbm.tudelft.nl/.

Figure 2 shows an example of a choice task from the experiment. Respondents were given the following instructions:

- Imagine you are cycling from your work, train station, school, or daily activity to your home.
- There are two cycling routes you can take.
- The routes differ only in **travel time**, the number of **traffic lights**, and the **cycling environment** shown in the photo.
- In all other aspects, such as the weather, the cycling routes are the same.
- **Which route would you choose?**
- You can assume the following:
  - You are cycling alone.
  - You are not in a hurry.
  - The photo gives a good idea of what the **entire route** looks like.
  - The travel time is the **total time** for the route, **including waiting at traffic lights.**
  - The weather is partly cloudy with no rain.

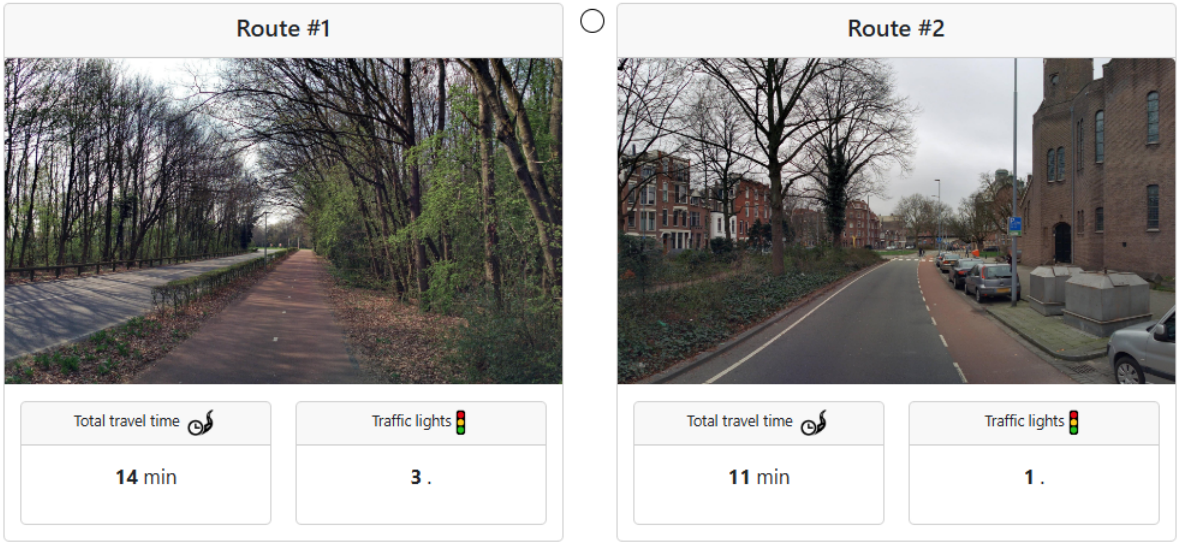


Figure 2: Example of a choice situation

2.2.4 Survey implementation

The survey was conducted between the end of September and the beginning of October 2024. Respondents were recruited through the panel company Cint (2024). The respondents were financially compensated and were individuals aged 18 years and above who engaged in cycling. To ensure that the sample was by and large representative of the Dutch population, entrance to the survey was controlled based on gender, age, and region through the data of CBS (2024). In total, 753 valid surveys have been collected. Regarding the images, a total of 6,484 unique images were used.

3 Methodology

3.1 Computer vision-enriched discrete choice models

In this study, we employ Computer Vision-enriched Discrete Choice Models (CV-DCMs) to estimate cyclists' route choice preferences (van Cranenburgh & Garrido-Valenzuela, 2025). CV-DCMs are grounded in Random Utility Theory (RUT) McFadden (1974), which assumes individuals select the maximum utility alternative. Furthermore, it assumes utility is composed of an observed part and an unobserved part from the researcher's perspective. Equation 1 shows the utility function of the linear-additive model, where i denotes the alternative and m denotes the attribute. Under the assumption that ε is i.i.d. Extreme Value distributed, the choice probabilities attain the well-known logit formula (2).

$$U_{in} = \underbrace{\sum_{m} \beta_m \cdot x_{imn}}_{\mathrm{V}_{in}} + \varepsilon_{in} \tag{1}$$

$$P_{in} = \frac{e^{V_{in}}}{\sum_j e^{V_{jn}}} \tag{2}$$

CV-DCMs combine traditional numerical attributes and information encoded in image data into the RUM-based discrete choice modelling framework, as shown in Figure 3. Images are represented as three-dimensional tensors, with pixels containing RGB colour channels. But, raw pixels cannot be used directly in the model. CV-DCMs use a feature extractor —typically a deep neural network— to extract the salient features from the image. These, in turn, are mapped linearly onto utility. Our application uses a Data Efficient Image Transformer (DeiT), a vision transformer, as the feature extractor (ImageNet, 2024; Touvron et al., 2021). This model produces a feature map of size 1 x 1000. Equation 3 shows its utility function.

$$U_{in} = \underbrace{\sum_{m} \beta_m \cdot x_{imn}}_{\text{Utility from numerical attributes}} + \underbrace{\sum_{k} \beta_k \cdot z_{ikn}}_{\text{Utility from image features}} + \varepsilon_{in} \tag{3}$$

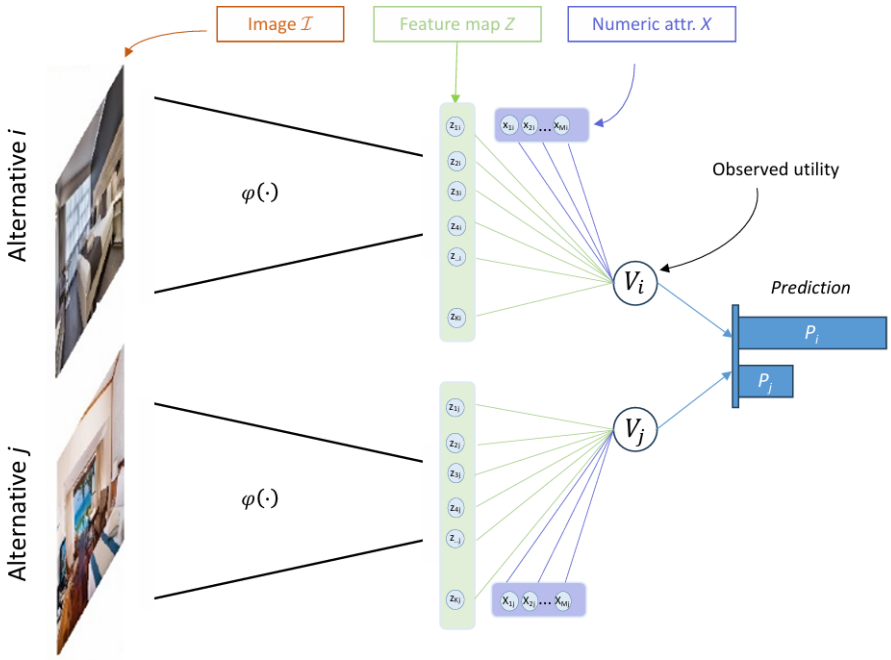


Figure 3: Model structure of the CV-DCM.
Adopted from van Cranenburgh and Garrido-Valenzuela, 2025

3.2 Training

3.2.1 Cross-entropy loss function

To train the CV-DCM, we use a cross-entropy loss function. Minimising the cross-entropy is equivalent to maximising the (log-)likelihood of the data. Furthermore, our loss function also includes an L2 regularisation term to prevent overfitting from happening. The loss function is shown in Equation 4, where β are the marginal utilities associated with the numerical attributes, and w are weights of the computer vision part of the model.

$$w, \beta = \arg\min_{w,\beta} \left(\underbrace{\frac{1}{N} \sum_{n=1}^{N} \sum_{j=1}^{J} y_{nj} \log(P_{nj} | X_{nj}, S_{jn}, \beta)}_{\text{Cross-entropy loss}} + \underbrace{\gamma \sum_{r=1}^{R} w_r^2}_{\text{L2 regularisation}} \right) \tag{4}$$

3.2.2 Training and hyperparameter tuning

We implemented and trained the CV-DCM using PyTorch. PyTorch is designed to train deep neural networks and facilitate GPU support, accelerating training. We set the hyperparameters to the optimal values found by (van Cranenburgh & Garrido-Valenzuela, 2025), which are reported in Table 4.

Table 4: Hyperparameters CV-DCM training

Hyperparameter	Value
Optimisation algorithm	Stochastic Gradient Descent
Learning Rate	1×10^{-5}
Batch Size	5
L2 weight decay (γ)	0.1

4 Results

4.1 Training and estimation results

Besides the CV-DCM (Model 1), we estimated two traditional discrete choice models for comparison. Model 2 is a linear-additive RUM-MNL model, which uses the travel time and the number of traffic lights as attributes only. Hence, this model ignores the presented images. Therefore, we expect this model's model fit to be lower than that of the CV-DCM. Model 3 extends Model 2 by also using cycling environment attributes of the place of the street-level image, which were merged onto data (see section 2.1). Table 5 reports the training and estimation results.

Table 5: Training and estimation results

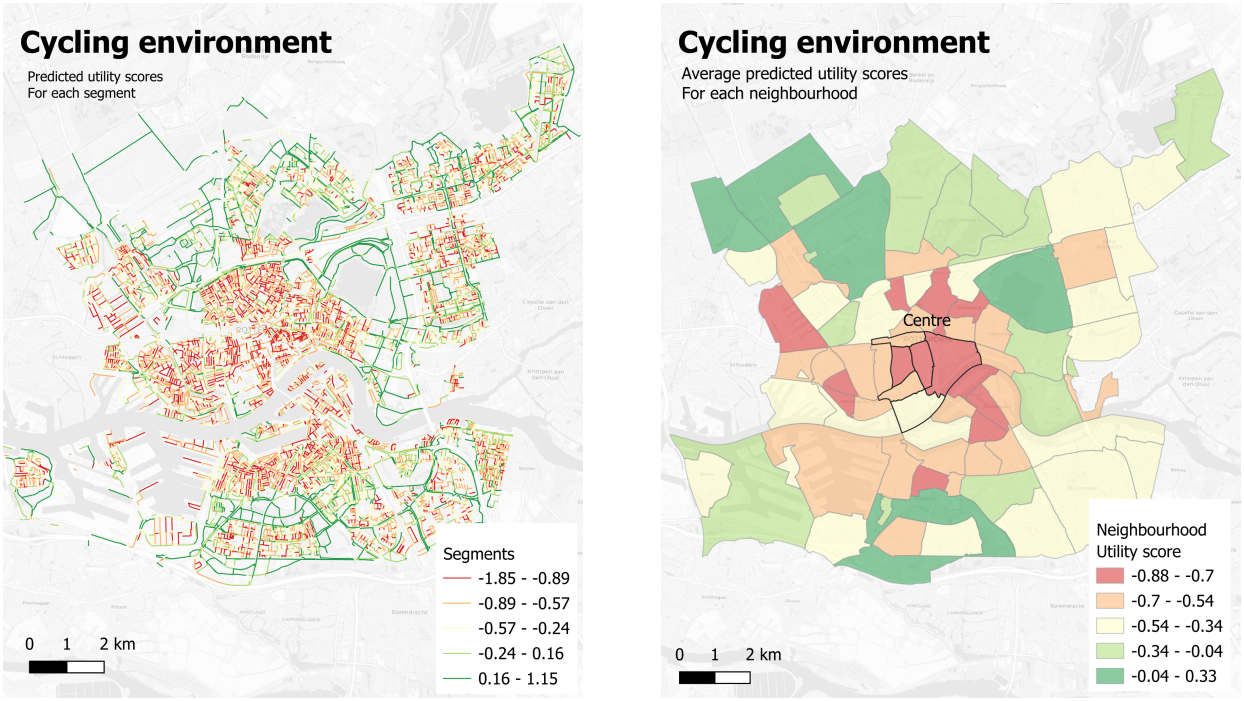
	Model 1			**Model 2**			**Model 3**		
Model type	CV-DCM			RUM-MNL			RUM-MNL		
Number of parameters	86m			2			15		
Estimation time	1.5 hour			<1 sec			<1 sec		
Train Set (N = 9,135)									
Log-Likelihood	-5,656			-6,157			-5,704		
Rho-squared	0.109			0.028			0.099		
Cross-entropy	0.617			0.674			0.624		
BIC	-			12,332			11,208		
Test Set (N = 2,128)									
Log-Likelihood	-1,326			-1,421			-1,339		
Rho-squared	0.101			0.036			0.092		
Cross-entropy	0.623			0.668			0.629		
BIC	-			2,857			2,793		
	est	**s.e.**	**p-val**	**est**	**s.e.**	**p-val**	**est**	**s.e.**	**p-val**
Parameters									
β_{tt}	-0.10			-0.11	0.012	0.00	-0.23	0.0139	0.00
β_{tl}	-0.24			-0.23	0.021	0.00	-0.31	0.0215	0.00
$\beta_{normalroad}$							0.00	Fixed	Fixed
$\beta_{suggestionlane}$							-0.05	0.054	0.37
$\beta_{separatedlane}$							0.36	0.054	0.00
$\beta_{klinkers}$							-0.11	0.041	0.01
β_{sunny}							0.09	0.033	0.01
$\beta_{brightness}$							0.04	0.008	0.00
β_{trees}							0.10	0.015	0.00
β_{water}							0.18	0.038	0.00
β_{house}							-0.31	0.049	0.00
$\beta_{industrial}$							-0.63	0.059	0.00
β_{gras}							0.24	0.038	0.00
$\beta_{parking}$							-0.09	0.039	0.02
β_{tram}							-0.23	0.107	0.03
β_{plants}							0.03	0.037	0.45
WTP traffic light [min]	2.38			2.12	0.2522	0.00	1.38	0.1034	0.00

We can make the following observations based on Table 5. Firstly, the Rho-square indicates that the CV-DCM model (Model 1) explains a reasonable amount of the variance in the choice data. In contrast, Model 2, which ignores the information encoded in the images, attains a very low model fit. From this comparison, we can infer that the computer vision part of the CV-DCM extracts meaningful information from the images to explain the route choice behaviour in the data. Similar to traditional DCMs, the β coefficients of the CV-DCM represent marginal utilities and can be used to calculate Willingness-to-Pay (WTP) estimates. Specifically, the WTP to reduce one traffic light can be determined by taking the ratio of the coefficients, β_{tl}/β_{tt}. Based on the CV-DCM, the WTP to eliminate a traffic light from the cycling route is estimated to be 2.38 minutes. Finally, examining Model 3, we observe that it achieves a significantly better fit than Model 2, though it still falls short of the CV-DCM. This

suggests that environmental features and the images share overlapping information. For example, the presence of a building along the cycling route can be detected both in the street-level image and through data extracted from a GIS database. Although it does not attain an as good fit as the CV-DCM, it provides more explanations because the betas associated with the cycling environment features are interpretable. In particular, Model 3 reveals that cyclists strongly prefer separated cycling lanes over normal roads. Specifically, the findings indicate that cyclists are willing to extend an 11-minute trip by 1.5 minutes to use a separate cycling lane rather than a mixed-traffic road. Similarly, natural features such as trees, water, and grass are significantly more appreciated than normal roads, whereas industrial areas, parking spaces, and trams are less favoured by cyclists.

4.2 Model application in Rotterdam

To investigate the spatial distribution of the attractiveness of the cycling environment in Rotterdam, we applied our trained CV-DCM model to 13k images taken across Rotterdam's cycling network. For each image, the CV-DCM predicts a utility level, which can be conceived as a measure of the cycling environment's attractiveness. Figures 4a and 4b show the results. In Figure 4a, the results are aggregated at the street-segment level, whereas in Figure 4b, the results are first aggregated at the neighbourhood level and subsequently grouped into five clusters using Jenks natural breaks (J. Chen et al., 2013). The legends display the corresponding colourmaps, but it is important to note that utility values do not have an absolute reference level (Train, 2009). Consequently, the mean utility is not zero-centred. Instead, the mean utility level is -0.46 in Figures 4a and 4b. Both figures highlight a noticeable pattern: the cycling environment tends to be relatively poor in the city centre but improves considerably in the periphery.



(a) Utility levels, aggregated at street segments

(b) Utility levels, aggregated at neighbourhood level

Figure 4: Spatial distribution of utility derived from the cycling environment

4.2.1 Explaining cycling environment utility levels

A drawback of the CV-DCM is that its computer vision part is opaque (van Cranenburgh & Garrido-Valenzuela, 2024). However, one can look at the most extreme predictions to elucidate what it has learned. Specifically, looking at the images that produce the lowest and highest utility levels provides insights into the factors the model uses for its predictions. Accordingly, Figure 5 shows nine random images from the highest and lowest utility score clusters used to produce Figure 4b. We observe that the highest utility levels are associated with separated cycling lanes, greenery and few buildings or traffic. In contrast, the lowest utility levels are associated with images of urban areas with buildings, vehicles, and shared narrow roads. This pattern aligns with prior findings that cyclists prefer routes with greenery and separated lanes (P. Chen et al., 2018; Rossetti et al., 2018).

(a) Images sampled from cluster 1

(b) Images sampled from cluster 5

Figure 5: Images with the highest (left) and lowest (right) utility levels

To further elucidate what the CV-DCM has learned about cyclists' preferences for cycling infrastructure, we grouped the images based on the type of lanes and the maximum car speed on the road. Figure 6 shows the results in a boxplot, in which the columns are sorted from lowest mean utility to highest mean utility. Figure 6 indicates that ordinary roads are the least preferred option, followed by cycling suggestion lanes. Interestingly, the maximum car speed (30 km/h or 50 km/h) appears to have little influence, possibly because participants did not infer speed from the images. Separated and solitary cycling lanes are the most preferred lanes.

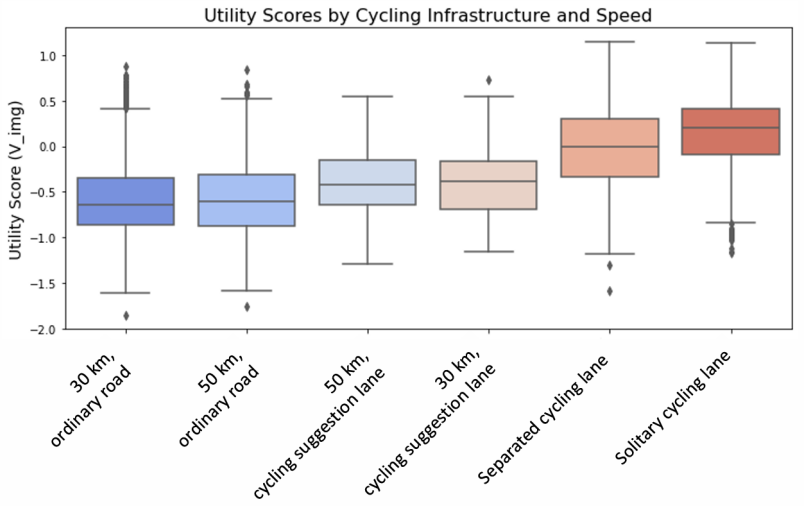


Figure 6: Utility scores by cycling infrastructure and Car speed

We grouped the images on the built environment categories, also used to create the stated choice experiment. Figure 7 shows that recreational areas, characterised by separated cycling lanes and greenery, are most preferred by cyclists. Main roads and neighbourhood access roads follow, while residential and industrial areas score similarly, with industrial areas slightly lower. Dense buildings and clinker roads likely reduce the appeal of residential areas.

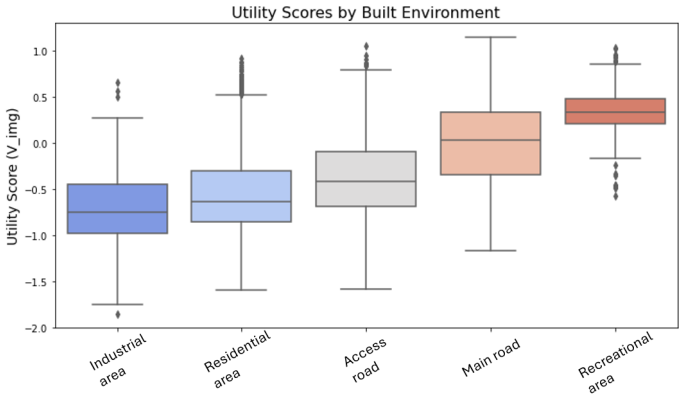


Figure 7: Utility scores by cycling infrastructure and Car speed

While this analysis highlights valuable insights, the correlation between cycling infrastructure and the built environment presents a challenge. For instance, recreational areas often feature solitary cycling lanes surrounded by greenery. This overlap complicates interpretation, as utility scores may reflect the combined appeal of infrastructure and environment rather than isolating their individual impact.

5 Conclusions and discussion

This paper investigated cyclists' preferences for cycling environments, such as lane type, pavement type, and greenery. To do so, we conducted a novel stated cycling route choice experiment in which respondents were presented alternatives that varied in terms of the cycling environment, travel time and the number of traffic lights. In contrast to previous studies, we used street-level images to represent the cycling environment. For the stated choice experiment, over 6k were used. We analysed our choice data using recently proposed model computer vision-enriched discrete choice models, which integrate computer vision into traditional discrete choice models and compared the results with traditional discrete choice models. Our results show that the cycling environment is an influential factor in cyclists' route choice behaviour. Specifically, we find strong support for the notion that cyclists prefer separated cycling lanes. Specifically, we find that cyclists are willing to detour 1.5 minutes to cycle on a separate lane instead of a mixed-traffic road for an 11-minute trip.

While the use of real-world images in stated choice experiments offers significant advantages, it also raises important questions. For example, it is plausible that the visual representation of the cycling environment disproportionately captures attention, potentially diverting focus from numerically presented attributes (e.g., travel time and traffic lights in our study) and ultimately leading to biased estimates. Future research could investigate this hypothesis through an A/B testing framework. Moreover, although images greatly enhance respondents' understanding of the choice context compared to text and numerical descriptions alone, the presented scenarios may still deviate from real-world conditions. In our case, single images are used to represent entire routes, whereas in reality, the cycling environment can vary significantly along the route. Future studies could address this limitation by incorporating a series of images or videos to provide a more realistic representation of the choice situation. Finally, while the parameters of the CV-DCM associated with numerical attributes retain their interpretability, the weights derived from the computer vision model component lack transparency. Future research could explore model structures that enhance the interpretability of the computer vision component.

References

Bernardi, S., La Paix-Puello, L., & Geurs, K. (2018). Modelling route choice of dutch cyclists using smartphone data. *Journal of Transport and Land Use*, *11*(1). https://doi.org/10.5198/jtlu.2018.1143

CBS. (2024). Cbs. https://opendata.cbs.nl/statline/#/CBS/nl/

Chen, J., Yang, S. T., Li, H. W., Zhang, B., & Lv, J. R. (2013). Research on geographical environment unit division based on the method of natural breaks (jenks). *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XL-4/W3*, 47–50. https://doi.org/10.5194/isprsarchives-XL-4-W3-47-2013

Chen, P., Shen, Q., & Childress, S. (2018). A gps data-based analysis of built environment influences on bicyclist route preferences. *International Journal of Sustainable Transportation*, *12*(3), 218–231. https://doi.org/10.1080/15568318.2017.1349222

ChoiceMetrics. (2024). Ngene. https://choice-metrics.com/download.html

Cint. (2024). Cint. https://www.cint.com/

Costa, M., Azevedo, C. L., Siebert, F. W., Marques, M., & Moura, F. (2024). Unraveling the relation between cycling accidents and built environment typologies: Capturing spatial heterogeneity through a latent class discrete outcome model. *Accident Analysis & Prevention*, *200*, 107533.

Costa, M., Marques, M., Azevedo, C. L., Siebert, F. W., & Moura, F. (2024). Which cycling environment appears safer? learning cycling safety perceptions from pairwise image comparisons. *arXiv preprint arXiv:2412.09835*.

Cyclomedia. (2023). Panoramafoto's rotterdam actueel. https://www.gis.rotterdam.nl/gisweb2/default.aspx

Dubey, A., Naik, N., Parikh, D., Raskar, R., & Hidalgo, C. A. (2016). Deep learning the city: Quantifying urban perception at a global scale. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Computer vision – eccv 2016* (pp. 196–212). Springer International Publishing.

Dutch Cycling Embassy. (2018). Dutch cycling vision. https://dutchcycling.nl/wp-content/uploads/2022/05/Dutch_Cycling_Vision_EN.pdf

European Union. (2023). The importance of active mobility in the face of climate change. https://www.euroclima.org/en/recent-events-urban/articles-and-interviews/1981-world-bike-day-the-importance-of-active-mobility-in-the-face-of-climate-change

Fietsersbond. (2020). Data fietsersbond. https://www.fietsersbond.nl/

Fosgerau, M., Łukawska, M., Paulsen, M., & Rasmussen, T. K. (2023). Bikeability and the induced demand for cycling. *Proceedings of the National Academy of Sciences*, *120*(16). https://doi.org/10.1073/pnas.2220515120

Gössling, S., & McRae, S. (2022). Subjectively safe cycling infrastructure: New insights for urban designs. *Journal of Transport Geography*, *101*, 103340. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2022.103340

Heinen, E., van Wee, B., & Maat, K. (2010). Commuting by bicycle: An overview of the literature. *Transport Reviews*, *30*(1), 59–96. https://doi.org/10.1080/01441640903187001

ImageNet. (2024). Imagenet. https://image-net.org/index.php

I&W. (2020). Klimaatbeleid. https://www.rijksoverheid.nl/onderwerpen/klimaatverandering/klimaatbeleid

Kaplan, S., & Prato, C. (2015). A spatial analysis of land use and network effects on frequency and severity of cyclist-motorist crashes in the copenhagen region. *Traffic Injury Prevention*, *16*(7), 724–731.

Liu, L., Silva, E. A., Wu, C., & Wang, H. (2017). A machine learning-based method for the large-scale evaluation of the qualities of the urban environment. *Computers, environment and urban systems*, *65*, 113–125.

Liu, Y., Yang, D., Timmermans, H. J., & de Vries, B. (2020). Analysis of the impact of street-scale built environment design near metro stations on pedestrian and cyclist road segment choice: A stated choice experiment. *Journal of transport geography*, *82*, 102570.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, *3*(4), 303–328. https://doi.org/https://doi.org/10.1016/0047-2727(74)90003-6

Pinker, S. (Ed.). (1988). *Visual cognition*. MIT Press. https://archive.org/details/visualcognition00pink

Plotly. (2024). Data apps for production. https://plotly.com/

Rose, J. M., & Bliemer, M. C. J. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews*, *29*(5), 587–617. https://doi.org/10.1080/01441640902827623

Rossetti, T., Guevara, A., Galilea, P., & Hurtubia, R. (2018). Modeling safety as a perceptual latent variable to assess cycling infrastructure. *Transportation Research Part A: Policy and Practice*, *111*, 252–265. https://doi.org/https://doi.org/10.1016/j.tra.2018.03.019

Significance. (2022). Values of time, reliability and comfort in the netherlands 2022. https://english.kimnet.nl/publications/publications/2024/04/03/new-values-of-travel-time-reliability-and-comfort-in-the-netherlands

Ton, D., Cats, O., Duives, D., & Hoogendoorn, S. (2017). How do people cycle in amsterdam? estimating cyclists' route choice determinants using gps data from an urban area. *Transportation Research Record: Journal of the Transportation Research Board, 2662*, 1–10. https://doi.org/10.3141/2662-09

Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., & Jégou, H. (2021). Training data-efficient image transformers distillation through attention. https://arxiv.org/abs/2012.12877

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

van Cranenburgh, S., & Garrido-Valenzuela, F. (2024). A utility-based spatial analysis of residential street-level conditions; a case study of rotterdam. *arXiv preprint arXiv:2410.17880*.

van Cranenburgh, S., & 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.

Verhoeven, H., Van Hecke, L., & Dyck, V. (2018). Differences in physical environmental characteristics between adolescents' actual and shortest cycling routes: A study using a google street view-based audit. *International Journal of Health Geographics*, *17*. https://doi.org/https://doi.org/10.1186/s12942-018-0136-x

Visual Crossing. (2024). Weather data. https://www.visualcrossing.com/

Wade, N. J., & Swanston, M. (2013). *Visual perception: An introduction* (3rd). Psychology Press. https://www.taylorfrancis.com/books/mono/10.4324/9780203082263/visual-perception-nicholas-wade-mike-swanston

Wang, Y., Chau, C. K., Ng, W. Y., & Leung, T. M. (2016). A review on the effects of physical built environment attributes on enhancing walking and cycling activity levels within residential neighborhoods. *Cities*, *50*, 1–15. https://doi.org/https://doi.org/10.1016/j.cities.2015.08.004

Yang, Y., Wu, X., Zhou, P., Gou, Z., & Lu, Y. (2019). Towards a cycling-friendly city: An updated review of the associations between built environment and cycling behaviors (2007–2017). *Journal of Transport Health*, *14*, 100613. https://doi.org/https://doi.org/10.1016/j.jth.2019.100613

Ye, Y., Zhong, C., & Suel, E. (2024). Unpacking the perceived cycling safety of road environment using street view imagery and cycle accident data. *Accident Analysis & Prevention*, *205*, 107677.

Zhang, F., Salazar-Miranda, A., Duarte, F., Vale, L., Hack, G., Chen, M., Liu, Y., Batty, M., & Ratti, C. (2024). Urban visual intelligence: Studying cities with artificial intelligence and street-level imagery. *Annals of the American Association of Geographers*, *114*(5), 876–897.

Zimmermann, M., Mai, T., & Frejinger, E. (2017). Bike route choice modeling using gps data without choice sets of paths. *Transportation Research Part C: Emerging Technologies*, *75*, 183–196. https://doi.org/https://doi.org/10.1016/j.trc.2016.12.009