DigitalFUTURES 2024

Investigating associations between built environments and cycling behaviour using street view imagery and Strava Metro data:
A case study in City of Sydney, Australia

Hongming Yan / University of Technology Sydney, Xiaoran Huang / North China University of Technology, Jiaxin Liu / Politecnico di Milano, Sumita Ghosh / University of Technology Sydney, Martin Bryant / University of Technology Sydney

2024.07.07

The 6th International Conference on Computational Design and Robotic Fabricaion (CDRF 2024)

1. Introduction

2. Research gap

- Current research primarily focuses on correlational analysis using data collected from interviews and questionnaires with data from the SVI, with less exploration given to the integration of ***real-time cycling*** data.

- The majority of data derived from SVI represents only a minor portion of the environmental factors selected for previous studies, suggesting that ***the use of SVI is not yet exhaustive***.

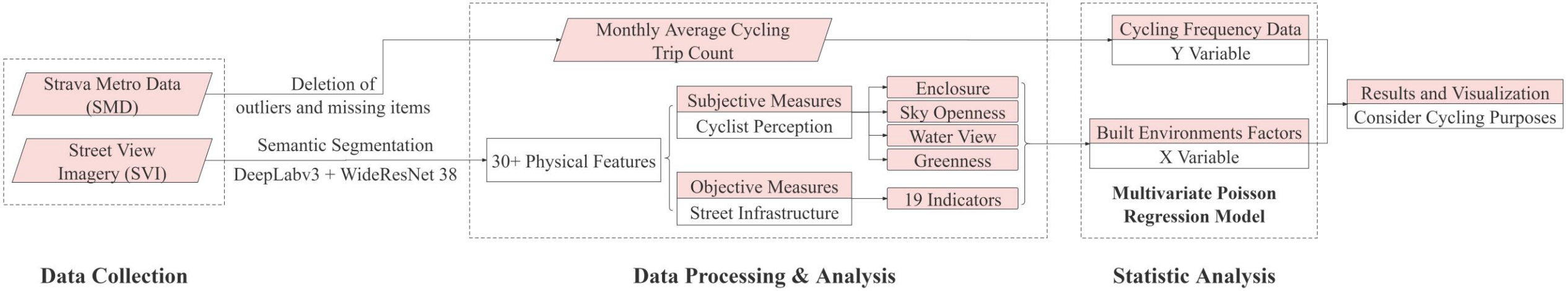
- While a few studies have utilized the Strava data to analyze urban cycling behavior, such research combining the active transport data with SVI to investigate the influence of built environment factors on cycling behavior is ***relatively rare in Australia***.

3. Research objectives

- This pilot study aims to validate the built environment factors affecting cycling behaviour in the central area of Sydney, Australia, by utilizing real-time cycling frequency data and a more comprehensive set of subjective and objective environmental factors obtained solely from SVI.
- The study further explored whether different built environment factors impact cyclist behaviour for different cycling purposes, by comparing the associations on commute cyclist and those on leisure cyclist.

4. Methodology

4.1 Research framework



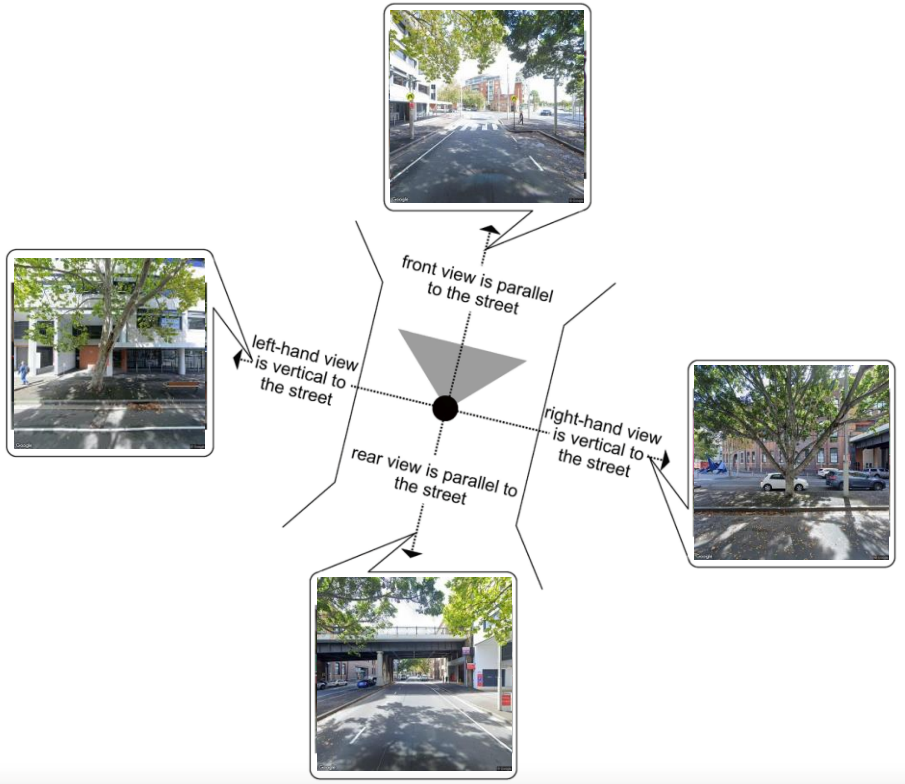
4. Methodology

4.2 Data collection – street view images

Google Maps

Street View

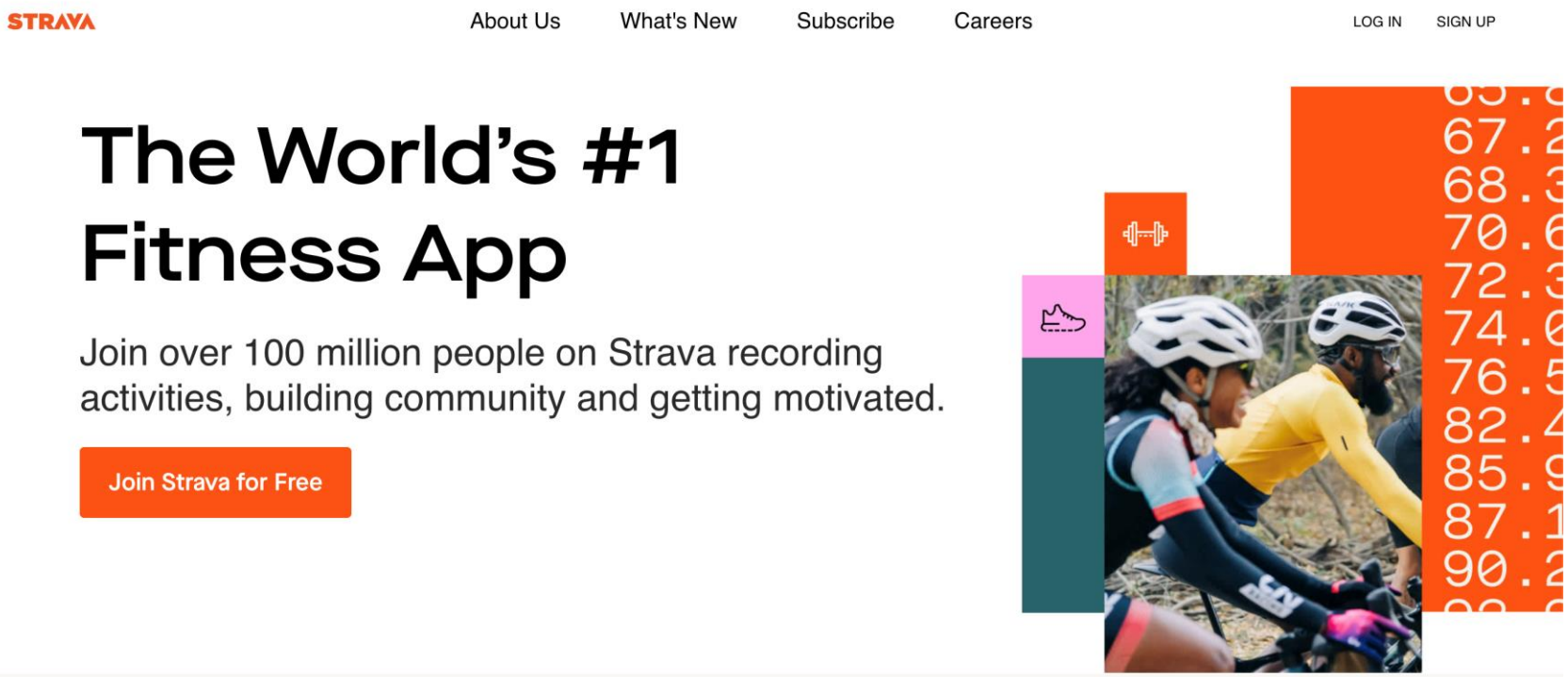
- 50m spacing
- Image size: 640 x 640
- Field of view: 90°
- 11618 synthetic panorama, 9301 images left after filtering within City of Sydney



Heading 0° Heading 90° Heading 180° Heading 270°

4. Methodology

4.3 Data collection – cycling frequency data



STRAVA | METRO

- Open source data
- Location: City of Sydney
- Time frame: Monthly
- Time period: January 2023 - December 2023

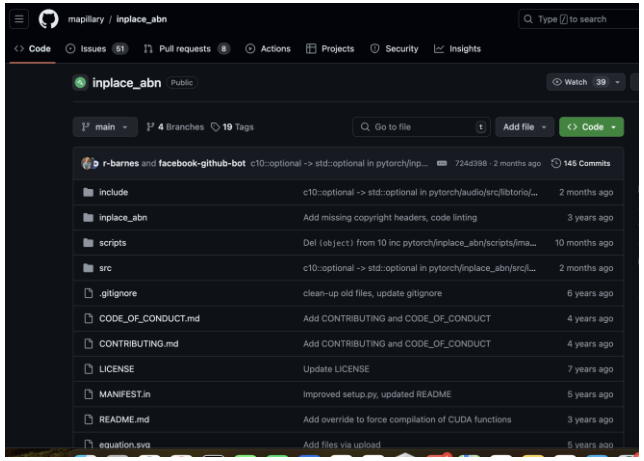
4. Methodology

4.4 Data processing – semantic segmentation (pre-trained Inplace_abn model Bulo et al. (2018))

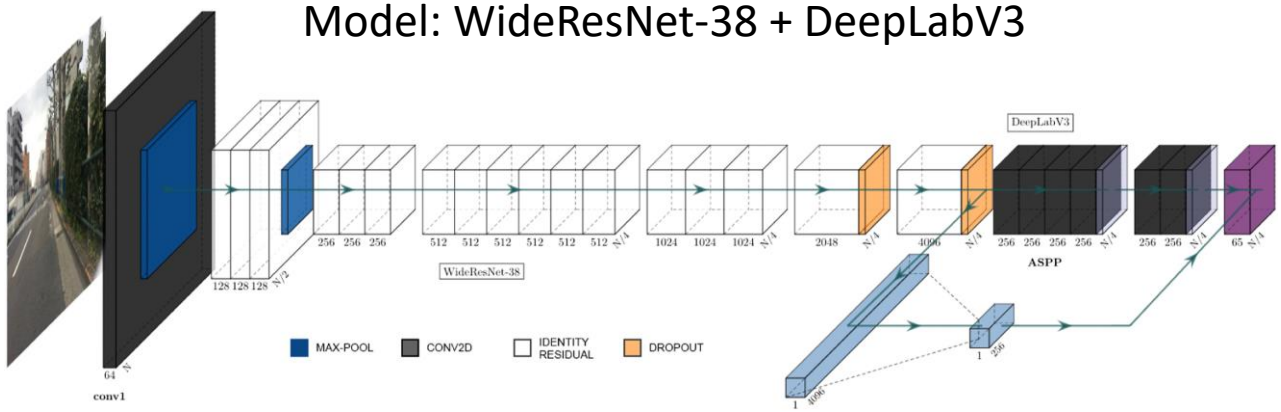
Dataset

Mapillary Vistas

Inplace_abn model



Model: WideResNet-38 + DeepLabV3



Source: Yap, W., Chang, J.-H., & Biljecki, F. (2023). Incorporating networks in semantic understanding of streetscapes: Contextualising active mobility decisions. *Environment and Planning B: Urban Analytics and City Science*, *50*(6), 1416–1437. https://doi.org/10.1177/23998083221138832

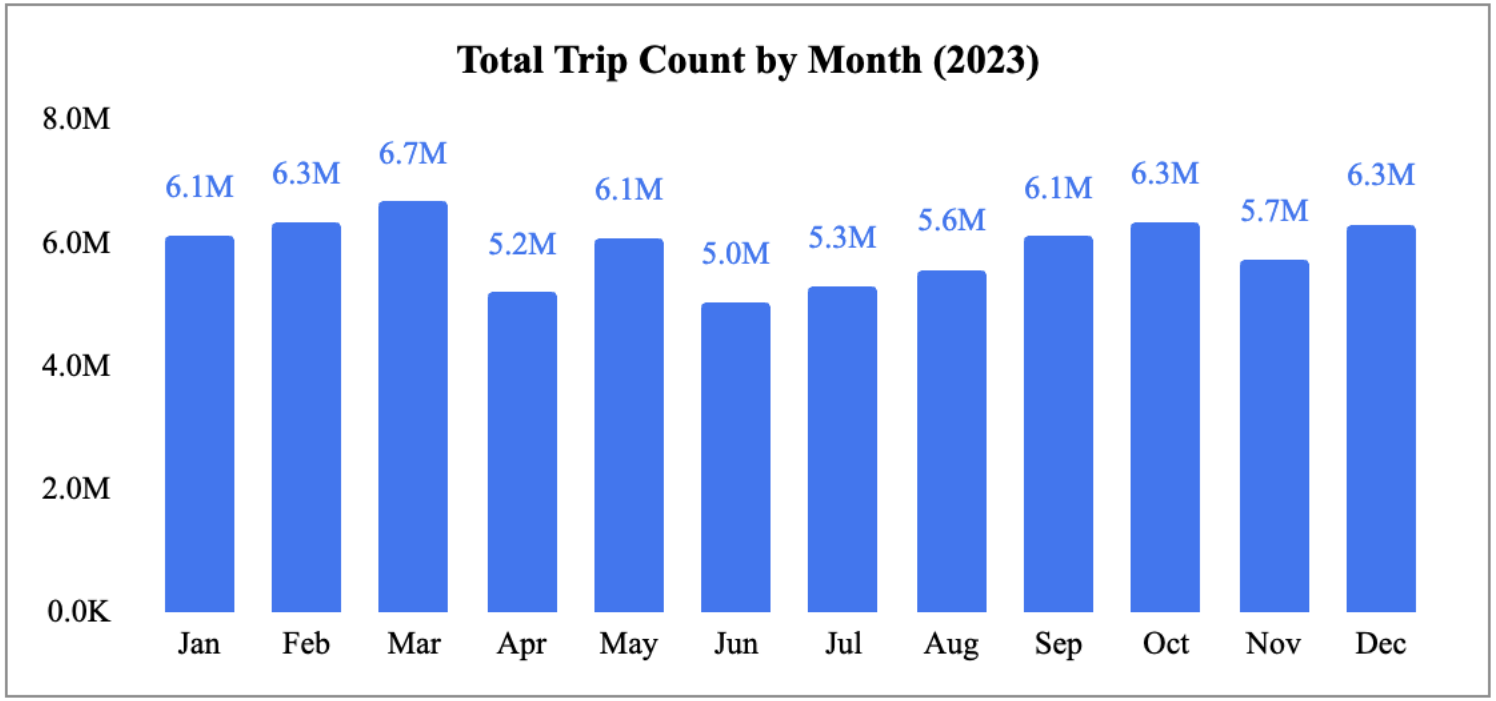
4. Methodology

4.5 Data processing – cycling frequency data: descriptive analysis

Strava Metro Data of City of Syndey 2023				
Total Segment Count			24,502	
Total Trip Count			71,143,145	
Commute Trip Count			24,391,075	(34%)
Leisure Trip Count			46,792,780	(66%)
Trip Count by Time	Morning	Midday	Evening	Overnight
	58%	17%	23%	3%
Cycling Participation by Gender	Male	Female	Unspecified	
	86%	12%	2%	
Cycling Participation by Age	18-34	35-54	55-64	65+
	32%	51%	14%	4%

4. Methodology

4.6 Data processing – cycling frequency data: descriptive analysis

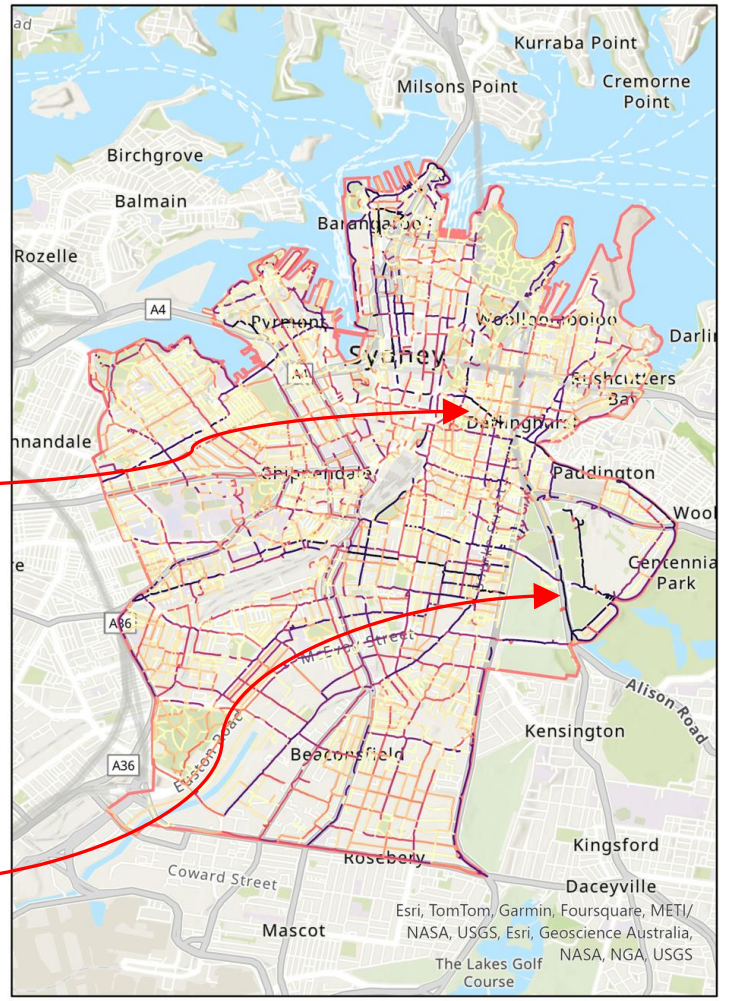


5. Results & Discussion

It is evident that the majority of higher frequency cycling routes are located on primary roads, while secondary and tertiary roads exhibit lower cycling frequencies.

Oxford Street

ANZAC Parade

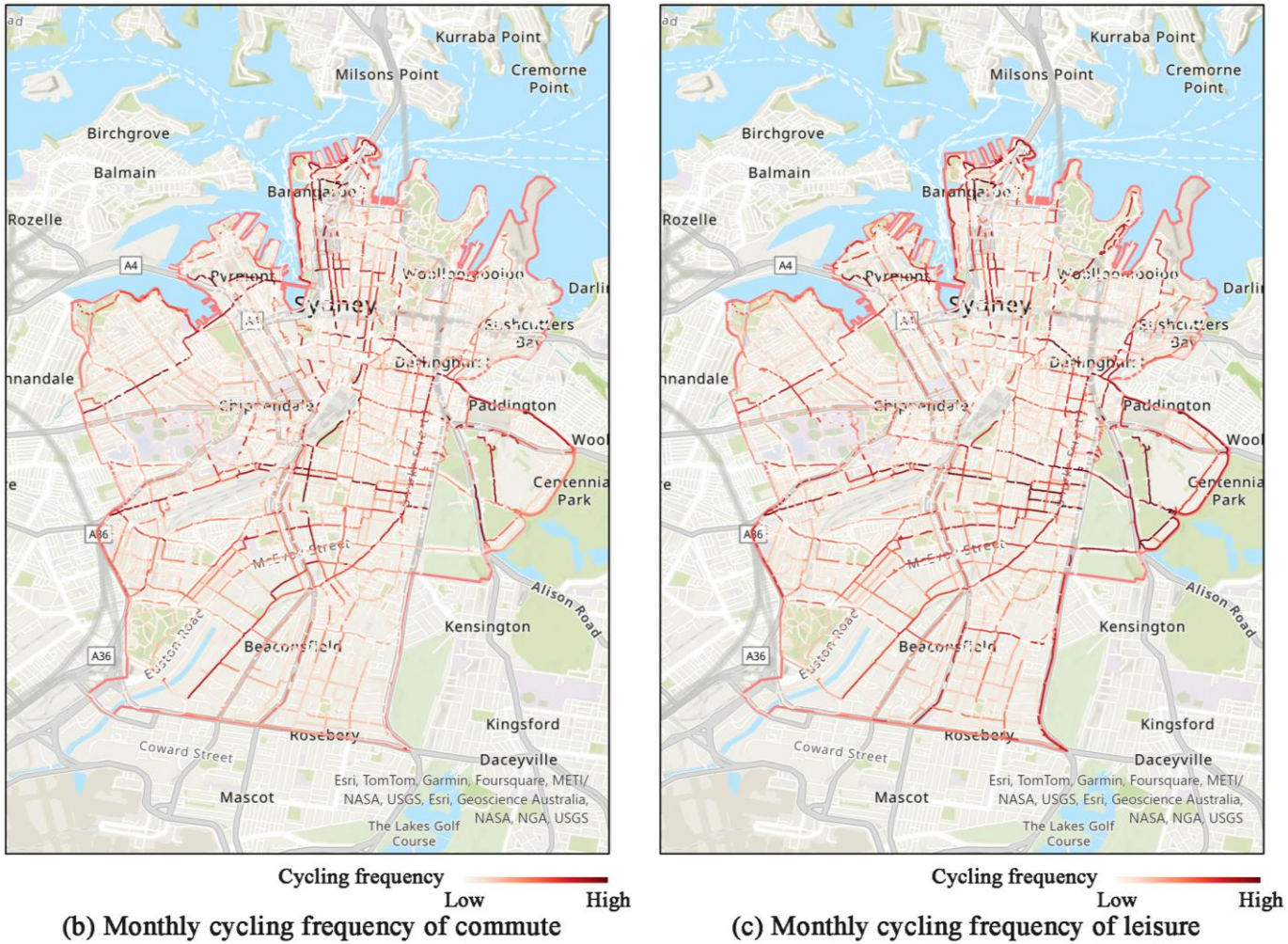


(a) Monthly cycling frequency of commute and leisure

5. Results & Discussion

- Cycling for commuting purposes is more frequent in **central urban areas, particularly near public transport stations**.

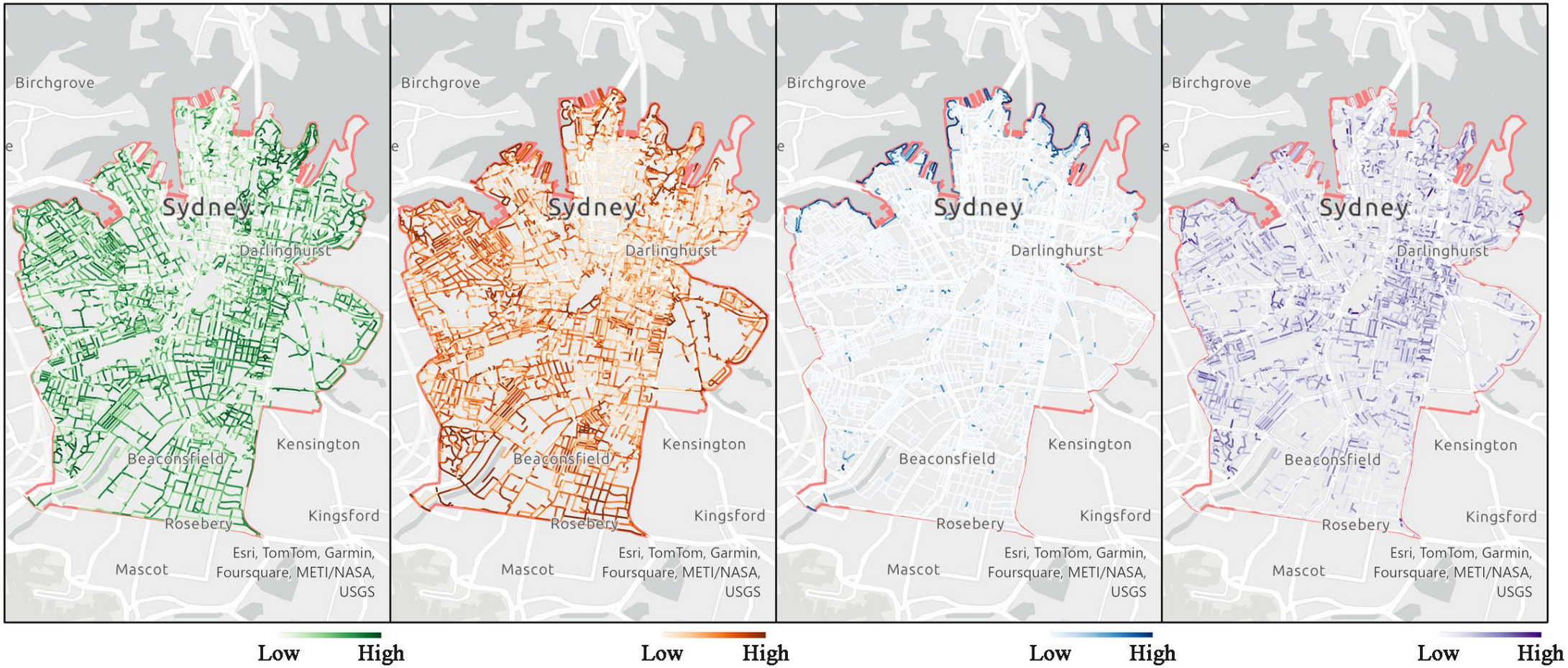
- Streets with higher cycling frequencies for leisure purposes are predominantly **located around urban green spaces and waterfront areas**.



(b) Monthly cycling frequency of commute

(c) Monthly cycling frequency of leisure

5. Results & Discussion



(a) Visual ratio of vegetation (b) Visual ratio of sky openness (c) Visual ratio of water view (d) Visual ratio of enclosure

5. Results & Discussion

Multivariate Poisson Regression

$$\ln[E(Y \mid X)] = \ln(t) + \beta_0 + \beta_1 X_1 + \beta_p X_p$$

Model predictor	Model 1: Total cycling frequency	Model 2: Commute cycling frequency	Model 3: Leisure cycling frequency
	Coef. (p-value)	Coef. (p-value)	Coef. (p-value)
Cycling perception indicators			
Vegetation	0.009 ***	0.006 ***	0.010 ***
Water	0.042 ***	-0.067 *	0.061 ***
Sky	-0.014 ***	-0.018 ***	-0.012 ***
Enclosure	-0.647 ***	-0.577 ***	-0.685 ***
Street infrastructure indicators (presence of ...)			
Bike.Lane.True	0.135 **	0.259 ***	0.073
Bike.Rack.True	0.142 *	0.146 .	0.142 *
Rail.Track.True	0.125 .	0.200 **	0.085
On.Street.Parking.True	0.199 ***	0.201 ***	0.200 ***
Sidewalk.True	0.095	0.060	0.104
Crosswalk.True	0.065	0.090 .	0.052
Curb.Cut.True	-0.097 **	-0.103 *	-0.096 **
Traffic.Light.True	0.350 ***	0.360 ***	0.347 ***
Traffic.Sign.True	-0.020	-0.002	-0.028
Traffic.Sign.Frame.True	-0.346 *	-0.499 **	-0.287 *
Street.Light.True	0.042	0.065	0.034
Junction.Box.True	-0.006	0.003	-0.011
Surveillance.True	-0.172	-0.177	-0.172
Pothole.True	-0.051	-0.151	-0.003
Manhole.Catch.Basin.Ture	-0.041	0.006	-0.064 .
Signage.Ads.True	0.271 ***	0.275 ***	0.274 ***
Banner.Ads.True	0.117 *	0.152 *	0.099 .
Street.Amenities.True	-0.071 *	-0.067 .	-0.073 *
Utility.Pole.True	0.420 *	0.774 **	0.298 .

Coef. = Coefficient; p-value: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5. Results & Discussion

Model predictor	Model 1: Total cycling frequency	Model 2: Commute cycling frequency	Model 3: Leisure cycling frequency
	Coef. (p-value)	Coef. (p-value)	Coef. (p-value)
Cycling perception indicators			
Vegetation	0.009 ***	0.006 ***	0.010 ***
Water	0.042 ***	-0.067 *	0.061 ***
Sky	-0.014 ***	-0.018 ***	-0.012 ***
Enclosure	-0.647 ***	-0.577 ***	-0.685 ***

- The visibility of vegetation positively influenced cycling frequencies regardless of the rider's purpose.
- Views of water were strongly positively correlated with leisure cycling in the City of Sydney
- A clear sky was consistently associated with lower frequencies of cycling.
- The enclosure aspect showed a negative relationship in cycling frequency

5. Results & Discussion

- The presence of bike lanes and traffic lights was positively correlated with cycling frequencies, particularly commute cyclists.
- Features such as curb cuts and large traffic sign frames had a negative effect on cycling frequency.
- Advertisements in signage and banner were positively associated with cycling frequencies.

Model predictor	Model 1: Total cycling frequency	Model 2: Commute cycling frequency	Model 3: Leisure cycling frequency
	Coef. (p-value)	Coef. (p-value)	Coef. (p-value)
Street infrastructure indicators (presence of ...)			
Bike.Lane.True	0.135 **	0.259 ***	0.073
Bike.Rack.True	0.142 *	0.146 .	0.142 *
Rail.Track.True	0.125 .	0.200 **	0.085
On.Street.Parking.True	0.199 ***	0.201 ***	0.200 ***
Sidewalk.True	0.095	0.060	0.104
Crosswalk.True	0.065	0.090	0.052
Curb.Cut.True	-0.097 **	-0.103 *	-0.096 **
Traffic.Light.True	0.350 ***	0.360 ***	0.347 ***
Traffic.Sign.True	-0.020	-0.002	-0.028
Traffic.Sign.Frame.True	-0.346 *	-0.499 **	-0.287 *
Street.Light.True	0.042	0.065	0.034
Junction.Box.True	-0.006	0.003	-0.011
Surveillance.True	-0.172	-0.177	-0.172
Pothole.True	-0.051	-0.151	-0.003
Manhole.Catch.Basin.Ture	-0.041	0.006	-0.064 .
Signage.Ads.True	0.271 ***	0.275 ***	0.274 ***
Banner.Ads.True	0.117 *	0.152 *	0.099 .
Street.Amenities.True	-0.071 *	-0.067 .	-0.073 *
Utility.Pole.True	0.420 *	0.774 **	0.298 .

Coef. = Coefficient; p-value: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

6. Limitation & Future research

- Our segmentation utilized ***pre-trained IPAB models*** on old Mapillary Vistas dataset, which have a limitation on categories (66 labels) extracted from images. The new Mapillary Vistas dataset, featuring 124 labels capable of capturing more comprehensive street information, along with a more powerful model like One-former, can be utilized for segmentation tasks.

- Some limitations on the cycling data such as the ***absence of personal and economic data*** about cyclists restricts the range of our statistical evaluation. Also, this study did not include some variables, such as urban density, road width, and land-use data. Future studies should consider incorporating these variables.

- ***Spatial heterogeneity and the non-linear relationship*** between built environments and cycling behaviour should also be taken into account in future research.

- As a pilot study of City of Sydney, future investigation could explore the comparison among ***other Australian cities.***