

# Computer vision and neural networks in assessing high-performance city design

Dr. Kerry A. Nice

Transport, Health, and Urban Design Research Lab, Melbourne School of Design  
University of Melbourne

Australia - Asia Dialogue for Urban Innovation: Smart Cities

May 8, 2023



# Transport, Health and Urban Design Research Lab



The aim of the Transport, Health and Urban Design Research Lab (THUD) is to explore the effects of physical and social systems on the health of populations.

---

The design of towns, cities, neighbourhoods, and social systems greatly influences the social, economic, environmental, and health outcomes of populations.

At THUD, we develop and apply novel methods using big data, artificial intelligence, and advanced analytics to provide new insights thereby influencing urban policies, globally.

# What does low carbon, healthy, productive urban design look like?



## Improved urban planning can play a role in reducing Australia's leading causes of death

Figure 2: Leading underlying causes of death in Australia by age group, 2011–2013

	1st	2nd	3rd	4th	5th
Age < 1	Other Perinatal & congenital	Other SIDS	Other Ill-defined causes	External Accidental threats to breathing	Other Selected metabolic disorders
Age 1–14	External Land transport accidents	Other Perinatal & congenital	Cancer Brain cancer	External Accidental poisoning	Other Cerebral palsy & related
Age 15–24	External Suicide	External Land transport accidents	External Accidental poisoning	External Assault	External Event of undetermined intent
Age 25–44	External Suicide	External Accidental poisoning	External Land transport accidents	Circulatory Coronary heart disease	Cancer Breast cancer
Age 45–64	Circulatory Coronary heart disease	Cancer Lung cancer	Cancer Breast cancer	Cancer Colorectal cancer	External Suicide
Age 65–74	Circulatory Coronary heart disease	Cancer Lung cancer	Respiratory COPD	Circulatory Cerebrovascular disease	Cancer Colorectal cancer
Age 75–84	Circulatory Coronary heart disease	Circulatory Cerebrovascular disease	Other Dementia & Alzheimer disease	Cancer Lung cancer	Respiratory COPD
Age 85–94	Circulatory Coronary heart disease	Other Dementia & Alzheimer disease	Circulatory Cerebrovascular disease	Respiratory COPD	Circulatory Heart failure
Age 95+	Circulatory Coronary heart disease	Other Dementia & Alzheimer disease	Circulatory Cerebrovascular disease	Circulatory Heart failure	Respiratory Influenza & pneumonia

## Urban design, transport, and health 2



## Land use, transport, and population health: estimating the health benefits of compact cities

Mark Stevenson, Jason Thompson, Thiago Hérick de Sá, Reid Ewing, Dinesh Mohan, Rod McClure, Ian Roberts, Geetam Tiwari, Billie Giles-Corti, Xiaoduan Sun, Mark Wallace, James Woodcock

Using a health impact assessment framework, we estimated the population health effects arising from alternative land-use and transport policy initiatives in six cities. Land-use changes were modelled to reflect a compact city in which land-use density and diversity were increased and distances to public transport were reduced to produce low motorised mobility, namely a modal shift from private motor vehicles to walking, cycling, and public transport. The modelled compact city scenario resulted in health gains for all cities (for diabetes, cardiovascular disease, and respiratory disease) with overall health gains of 420–826 disability-adjusted life-years (DALYs) per 100 000 population. However, for moderate to highly motorised cities, such as Melbourne, London, and Boston, the compact city scenario predicted a small increase in road trauma for cyclists and pedestrians (health loss of between 34 and 41 DALYs per 100 000 population). The findings suggest that government policies need to actively pursue land-use elements—particularly a focus towards compact cities—that support a modal shift away from private motor vehicles towards walking, cycling, and low-emission public transport. At the same time, these policies need to ensure the provision of safe walking and cycling infrastructure. The findings highlight the opportunities for policy makers to positively influence the overall health of city populations.

Published Online  
September 23, 2016  
[http://dx.doi.org/10.1016/S0140-6736\(16\)30067-8](http://dx.doi.org/10.1016/S0140-6736(16)30067-8)

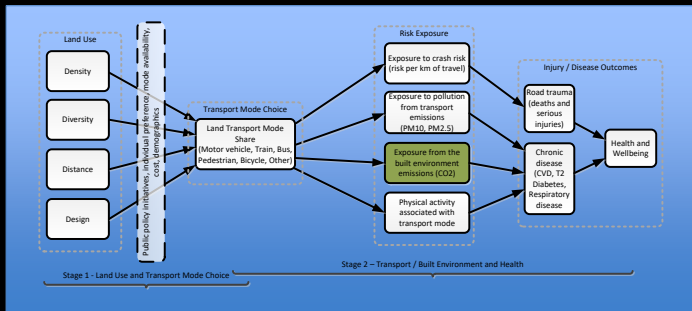
This is the second in a *Series of three papers about urban design, transport, and health*

University of Melbourne,  
Melbourne, VIC, Australia  
(Prof M Stevenson PhD,  
J Thompson PhD,  
Prof B Giles-Corti PhD);  
University of São Paulo, São  
Paulo, Brazil (T H de Sá PhD);  
University of Utah,  
Salt Lake City, UT, USA  
(Prof D Ewing PhD) - *Indian*

Stevenson, M., Thompson, J., Sá, T. H. D., Ewing, R., Mohan, D., McClure, R., Roberts, I., Tiwari, G., Giles-Corti, B., Sun, X., Wallace, M., and Woodcock, J. (2016). *Land use, transport, and population health: estimating the health benefits of compact cities*. *The Lancet*, 6736(16):1–11

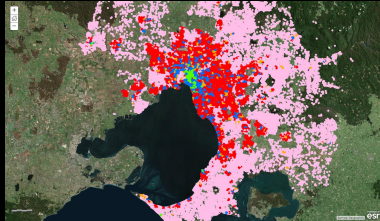
## Basic model

- Emerging from an existing 'deterministic', static model published in 'The Lancet', 2016



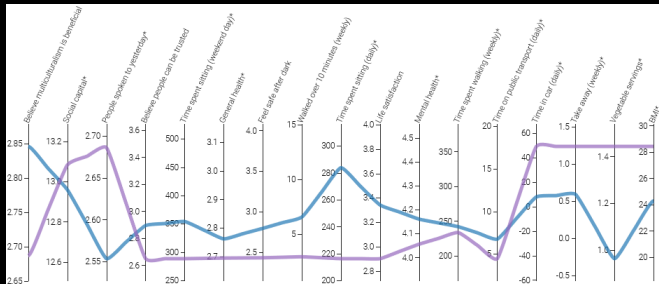
## Five Main Land-Use Clusters

1. High Density Inner City
2. Connected Inner Ring
3. Connected Pockets
4. Middle Suburbia, and
5. Urban Fringe



## Health Outcome Differences Between Clusters

1. High Density Inner City
2. Connected Inner Ring
3. Connected Pockets
4. Middle Suburbia, and
5. Urban Fringe





# City typology clustering

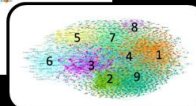
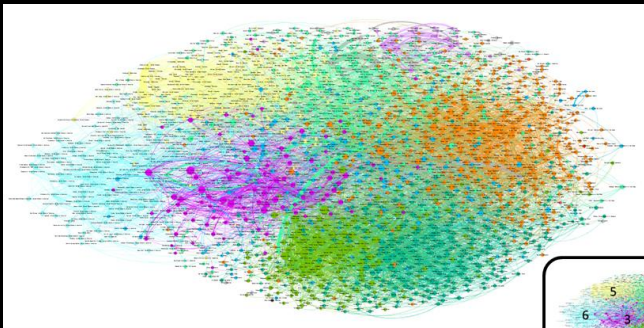
How to do global studies without global datasets? Use a neural network's confusion to identify cities by their maps (confusion = similar)

Each instance of confusion created a link between two cities, generating city 'clusters'












Road networks, public transport, open space, density, water, topography

## Confusion matrix SNA map



@agent\_iase

## 9 major city types

City Type	Name of City Type	Representative City	Sample Map Images
1	Informal (INF)	Hyderabad, India	
2	Irregular (IRR)	Beijing, China	
3	Large block (LBL)	Saint Petersburg, Russia	
4	Cul de sac (CUL)	Jakarta, Indonesia	
5	High transit (HTR)	Paris, France	
6	Motor city (MOT)	Los Angeles, USA	
7	Chequerboard (CHQ)	Mexico City, Mexico	
8	Intense (INT)	Osaka, Japan	
9	Sparse (SPR)	Dar Es Salaam, Tanzania	

## Geographic distribution



# City design relationship to road transport injury

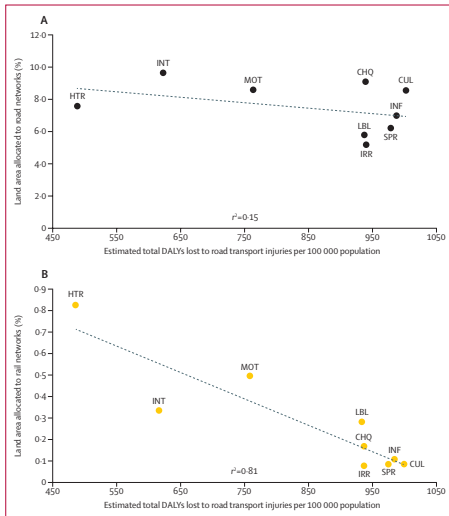
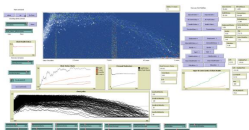
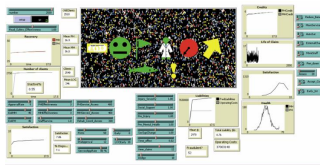
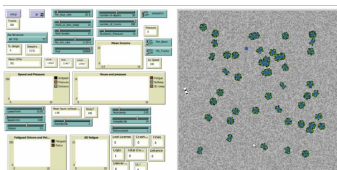
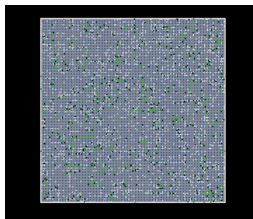


Figure 5: Association between city types dedicated to road (A) and rail (B) networks and DALYs lost to road transport injury per 100 000 population

# Agent-based modelling



# Modelling of SARS-CoV-2

## Modelling Victoria's escape from COVID-19

Table 1. Example parameter estimates and 'agent' characteristics.

Key Parameters	Parameter Estimates (varies under difference stage conditions – example parameter estimates only)
Physical distancing (% of people limiting movement and maintaining a distance of 1.5m in public) <sup>33</sup>	$m = 0.90\%$ , $sd = +/- 3\%$
Physical distancing - time (% of time that people successfully maintain a distance of 1.5m) <sup>33</sup>	$m = 0.90\%$ , $sd = +/- 3\%$
Proportion of essential workers <sup>4</sup>	$m = 25-100\%$ of working age-people (age 18-70)
Mean incubation period (days, log-normal) <sup>31</sup>	$m = 5.1$ , $sd = 1.5$
Mean illness period (days, log-normal) <sup>32</sup>	$m = 20.8$ , $sd = 2$
Mean adherence with isolation of infected cases <sup>5</sup>	$m = 93.3\%$ (beta distribution 28, 2; median = 94.3%, $sd = 4.5\%$ )
Likelihood of being traced by the health system per day after initial infection	25% – 80% of cases traced by day 5/6
Number of days after initial infection that new cases are reported <sup>6</sup>	6 <sup>4</sup>
Date of case simulation initialisation (Day 0)	September 1 <sup>st</sup> , 2020
Asymptomatic cases (% of cases) <sup>7,8</sup>	$m = 25\%$ , $sd = 3\%$
Infectiousness of asymptomatic cases vs symptomatic cases (per contact) <sup>38</sup>	$m = 33\%$ , $sd = 6\%$
Schools returns policies	True / False
Proportion of people wearing face-masks during interactions outside the home	50-90%
Reduction in transmission risk per contact for people wearing face-masks <sup>39</sup>	-25% (beta distribution)
Seeded cases for model initialisation based on prior 14 day average	-2400 on day 1 and 90 cases per day for 7 days
COVID-Safe App Uptake	$m = 30\%$ , $sd = +/- 3\%$
Agent Characteristics	Definition
Infection status	Infected, susceptible, recovered, deceased
Time now	The number of days (integer) since an infected person first became infected with SARS-CoV-2
Age-range	The age-bracket (categorical) of the person, set to Australian census data deciles from 0 to 100. Used in this simulation to capture differences in exposure risk through school closures and workforce status.
Risk of death	The overall risk of death (float) for each person based on their age-profile. Purely used in this simulation to remove the agents dying during the 100 day simulation time.
Location	Agents interact in over a 2-dimensional plane with their location recorded at each time-step via an x/y coordinate system (categorical).
Span	The distance the person moves around the environment away from their home location - longer distances result in higher likelihood of

Heading / Distance	close contact with novel other people (agents) in the model. The direction and extent of travel of the person at the current time-step. The heading and speed variables combine to create local communities and control interaction between and across communities. At higher lockdown stringency levels, agents are restricted to movement in areas closer to their home location.
Contacts	A count (integer) of contacts the person (agent) had interacted with in the past day as they moved within the model's environment. This is used in estimation of contacts with transmission potential each time-step and calculation of individual reproduction numbers at the end of infectious periods.

<sup>1</sup> Assumed parameter based on expert opinion in conjunction with available public face-coverage such as Google COVID-19 mobility reports. W 30% of the population potentially transmit infections widely through occasional travel to random locations.

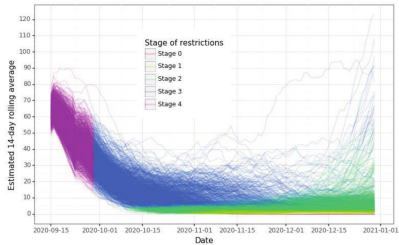
<sup>2</sup> YYY hospital based on recent Victorian estimates

<sup>4</sup> This reports all cases known to the model used on day 6 of their infection. In alternative modes, mode also allows for under-reporting under extreme pressure on the track and trace system (e.g., in uninvestigated scenarios).

<sup>6</sup> Contact tracing is fixed part of scenario, therefore no uncertainty.

# Modelling SARS-CoV-2

Figure 7: Estimated outcomes from 1,000 model runs from mid-September to 25 October, 2020. Each colour represents different stages as set out in the Victorian Roadmap.<sup>11</sup>



## Reopening too soon risks more lock-downs by Christmas

The fewer cases of coronavirus in the community when we ease, the lower the chance of locking down by Christmas.

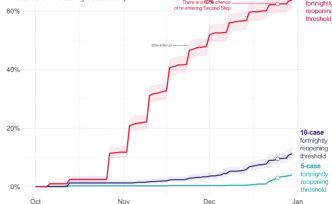
University of Melbourne modelled several policy scenarios.

If we ease restrictions when the average number of cases over the previous fortnight is 25 (350 cases total) then it's more likely than not that cases will get out of hand and restrictions will have to be reinstated to regain control and protect the health system.

Waiting until the average is 5 cases a fortnight – or 70 cases total – reduces the chance of increased restrictions before Christmas to just 3 in 100.

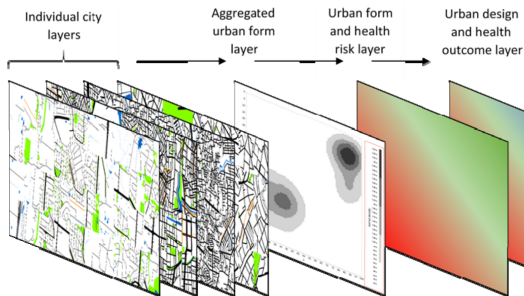
### Opening too early risks Second Step before Christmas

% chance of re-entering Second Step



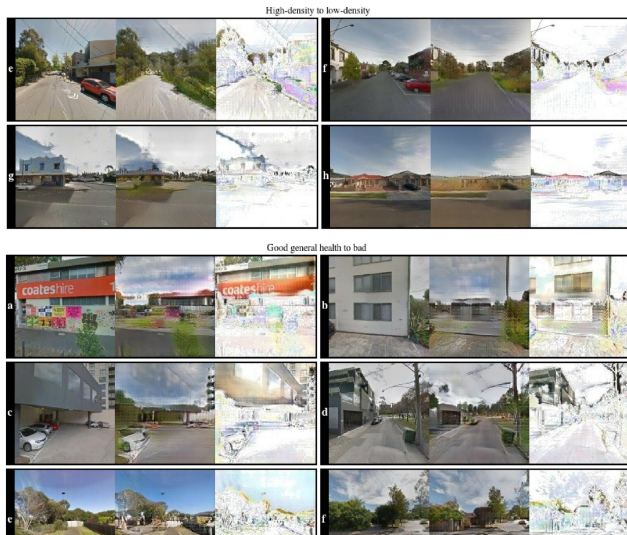


# Using computer vision to assess public health impacts of urban design



**Figure 2.** Representation of the city fingerprint and health impact process taking data from individual cities, combining all urban form into an aggregated 'fingerprint' map, and associating it with health risk and outcome data.

# GANs for healthy, safe urban design



# Using GANs to examine health impact of urban design



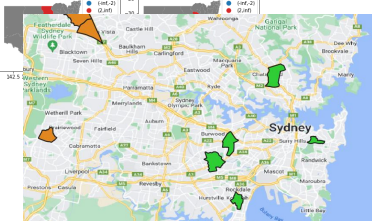
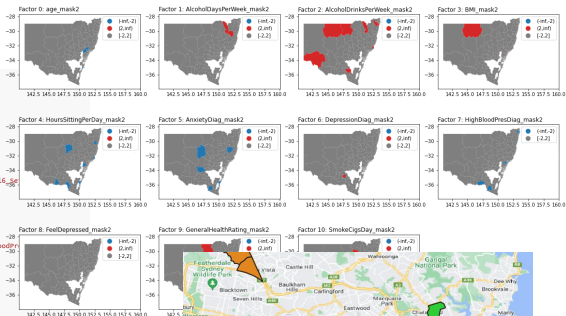
## New Images

```
for i in range(len(factors)):
    col = f'{factors[i]}_mask{z_value}'
    subfolder = 'Zscore_gt2' if z == 2 else 'zscore_lt2'
    sa2s = df_z2_numeric[df_z2_numeric[col]==0]['SA2_SDI616'].to_numpy()
    count = sa2s.size # shape[0]
    array = ' '.join(map(str, sa2s))
    asset = set(sa2s.flatten())
    data.append([factors[i], count, sa2s, asset, len(asset)])

df_z2_set = pd.DataFrame(data, columns=['Factor', 'Count', 'SA2_SDI616', 'SA2_SDI616_Set'])
df_z2_set.set_index('Factor', inplace=True)
display(df_z2_set)

Set_Age = df_z2_set.loc['age', 'SA2_SDI616_Set']
print(f'Set_Age: {Set_Age}')
Cat_Others = [i for i in factors if i not in ['age', 'AlcoholDaysPerWeek', 'HighBloodPressure']]
print(f'Cat_Others: {Cat_Others}')
Set_Others = [df_z2_set.loc[i, 'SA2_SDI616_Set'] for i in Cat_Others]
Set_Others_Union = set().union(*Set_Others)
Set_diff = Set_Age - Set_Others_Union #set A difference set B for (A - B)
print(f'Set_diff: {Set_diff}')

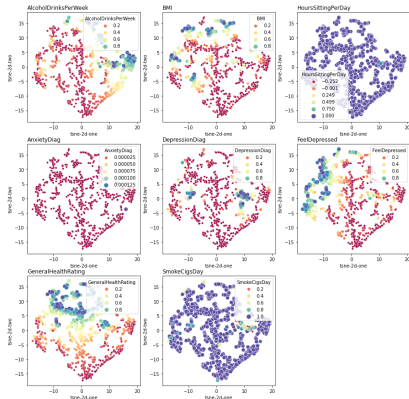
"""result
Set_Age: {11234, 11363, 11398, 11464, 11113, 11514, 11309, 11310, 11470, 11281, 11378, 11347, 11034,
11390}
Cat_Others: ['AlcoholDrinksPerWeek', 'BMI', 'HoursSittingPerDay', 'AnxietyDiag', 'DepressionDiag',
'FreeDepressed', 'GeneralHealthRating', 'SmokeCigsDay']
Set_diff: {11234, 11398, 11113, 11470, 11281, 11034, 11390}
"""
```



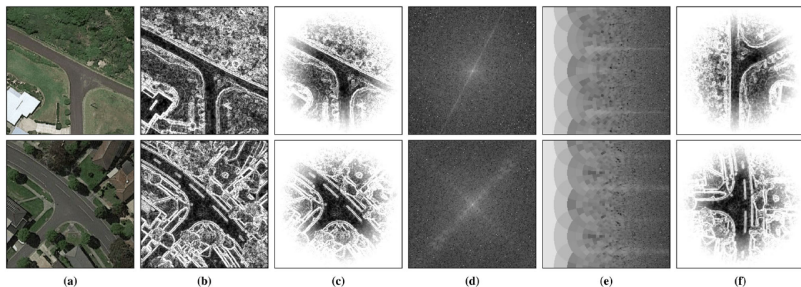
# Using GANs to examine health impact of urban design



t-SNE



# Identifying safe intersection design through satellite imagery and a deep autoencoder



**FIGURE 2** Illustration of preprocessing steps for two sample images. The plots show (a) the original Google Street View image, (b) line detection based on the Scharr operator, (c) emphasizing intersection center, (d) Fourier transform, (e) transformation to log-polar space to determine rotation angle, and (f) the final rotated image

# Identifying safe intersection design

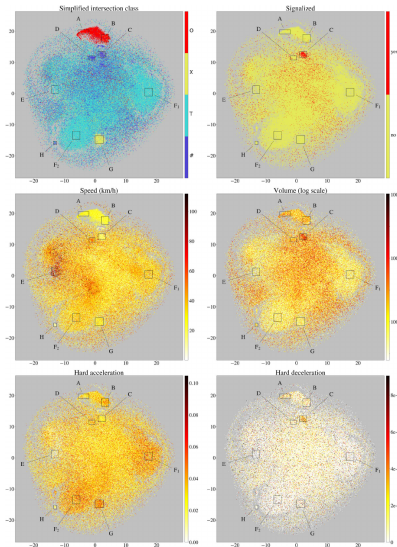


FIGURE 5 t-SNE clusters with matched OSM and telematics data (best viewed in color)

# Identifying safe intersection design



(a) Compact roundabout with stop line and ending bicycle lane (region B)



(b) Large roundabout with high traffic volume (region A)



(c) Parking spots along T-intersection (region F)



(d) Complex intersection without traffic lights (region C)



(e) Intersection with bus-only lane no traffic lights (region E)

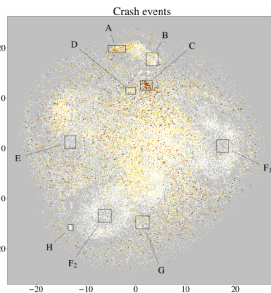



FIGURE 9 t-SNE clusters with matched crash events

## Real-time monitoring of driver drowsiness on mobile platforms using 3D neural networks

Jasper S. Wijnands<sup>1</sup>  · Jason Thompson<sup>1</sup> · Kerry A. Nice<sup>1</sup> · Gideon D. P. A. Aschwanden<sup>1</sup> · Mark Stevenson<sup>1,2,3</sup>

9736

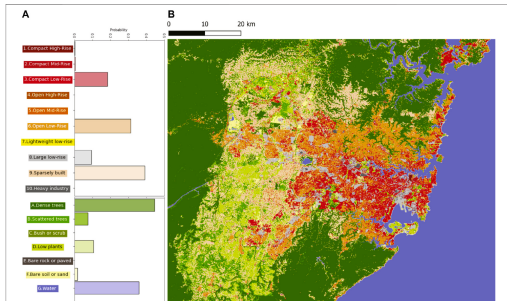
Neural Computing and Applications (2020) 32:9731–9743

**Fig. 2** Random distortions of a 10-frame sequence; only frames 1, 2 and 10 are displayed. The top row shows the original sequence, while rows A, B and C are different randomly generated pre-processed samples based on the original sequence

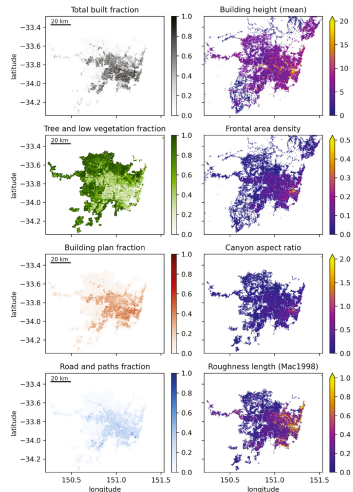




# Climate Infomatics: New sources of urban climate modelling data



**FIGURE 2** | Local Climate Zone map of greater Sydney, Australia **(B)** obtained using local training areas and LCZ Generator tool (Demuzere et al., 2021) shown together with the histogram distribution of LCZ classes across the area **(A)**.



# Detection of urban features for climate modelling



Contents lists available at ScienceDirect

Urban Climate

Journal homepage: [www.elsevier.com/locate/ucclim](http://www.elsevier.com/locate/ucclim)



## Sky pixel detection in outdoor imagery using an adaptive algorithm and machine learning



Kerry A. Nice<sup>a,\*,1</sup>, Jasper S. Wijnands<sup>a</sup>, Ariane Middel<sup>f</sup>, Jingcheng Wang<sup>h</sup>,  
Yiming Qiu<sup>g</sup>, Nan Zhao<sup>g</sup>, Jason Thompson<sup>g</sup>, Gideon D.P.A. Aschwanden<sup>g</sup>,  
Haifeng Zhao<sup>g</sup>, Mark Stevenson<sup>g,c</sup>

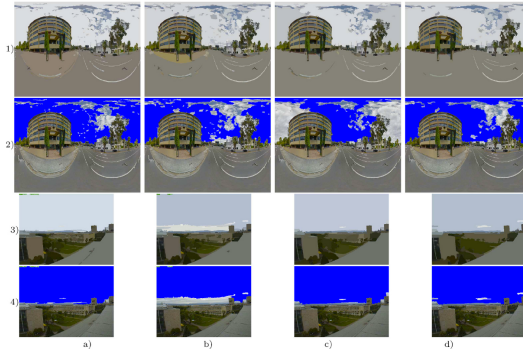
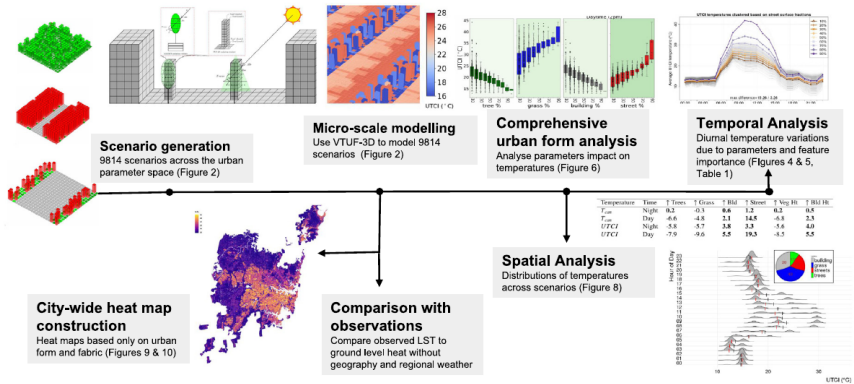


Fig. 4. Comparison outputs of intermediate mean shift segmentation algorithm processing steps using varying parameters, showing columns a) Mean\_3\_6\_100, b) Mean\_7\_6\_100, c) Mean\_5\_7\_210, d) Mean\_7\_8\_300 and intermediate mean shifted (GSV row 1, Skyfinder row 3) and the final marked images (GSV row 2, Skyfinder row 4).

# Impact of urban design on urban heat

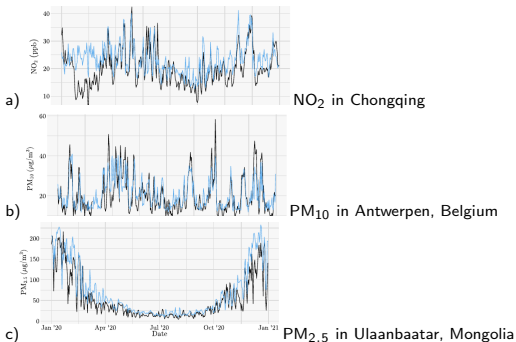
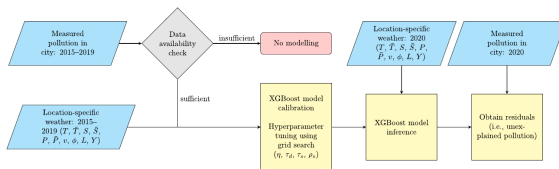


Nice, K. A., Nazarian, N., Lipson, M. J., Hart, M. A., Seneviratne, S., Thompson, J., Naserikia, M., Godic, B., and Stevenson, M. (2022). Isolating the impacts of urban form and fabric from geography on urban heat and human thermal comfort.

Building and Environment

# The impact of COVID-19 on global pollution

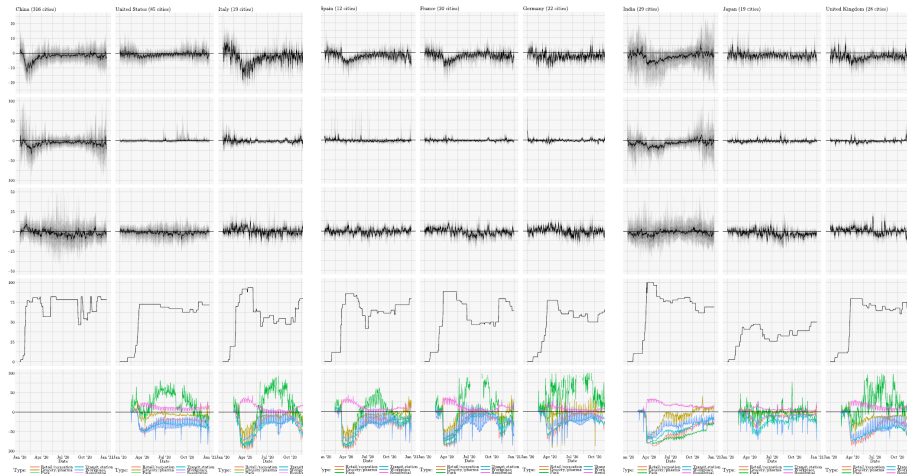
Combining ERA5 weather data with ground level pollution data, creating pollutant and city specific XGBoost model for 700 cities. Calculation of 2020 pollution anomalies in the absence of the COVID-19 pandemic.



Actual (black) and forecasted (blue) pollution levels

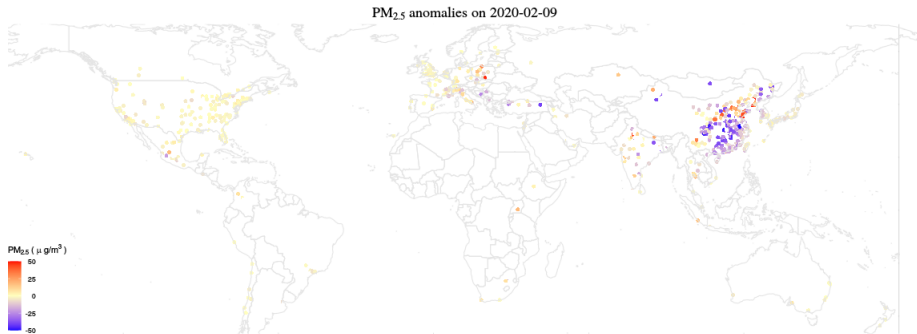
- Large February NO<sub>2</sub> reductions in China resulting from lockdowns.
- Aug 2020 Belgium heatwave accurately reflected through increased PM<sub>10</sub> levels.
- Temperature influence on pollution: seasonal trends of PM<sub>2.5</sub> in Mongolia from winter heating.

# Unexplained $\text{NO}_2$ , $\text{PM}_{2.5}$ , and $\text{O}_3$ across cities, stringency of COVID-19 restrictions, and mobility patterns from Google mobility data.



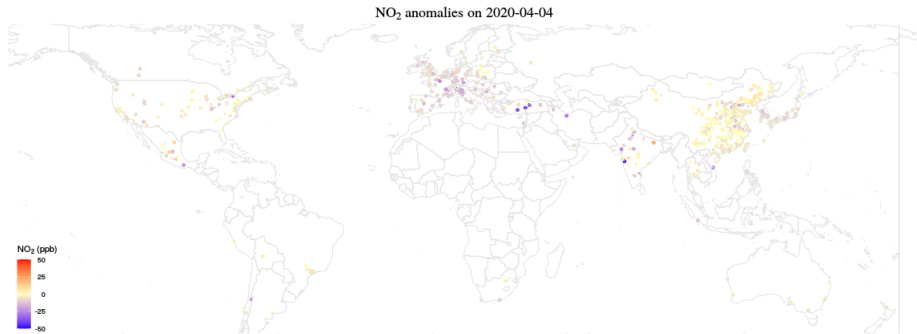
Large reductions in  $\text{NO}_2$  in most places across the first few months of 2020, largely returning to normal levels afterwards.  $\text{PM}_{2.5}$  reductions in many places across 2020.  $\text{O}_3$  increases in the first half of 2020 (due to  $\text{NO}_2$  reductions) but below normal levels mid-year.

# Global anomalies: PM<sub>2.5</sub> on 9 February 2020



Reductions of PM<sub>2.5</sub> in China in early February 2020.

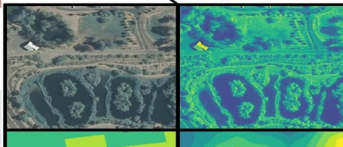
# Global anomalies: NO<sub>2</sub> on 4 April 2020



Reductions of NO<sub>2</sub> in Europe in early April. China had rebounded to normal levels.

## Imagery as a spatial data source

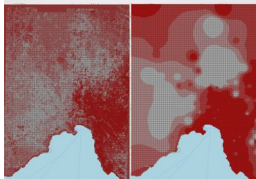
- Obtaining timely and accurate spatial data at scale is challenging
- Satellite and other imagery sources are rich in information but need to be processed
- Many advantages to using imagery
  - Updated often
  - Different types of imagery available: satellite, elevation maps, maps
  - Wide coverage – most urban areas
  - Ability to query different spatial characteristics
  - Customizable analysis
  - Reusable urban representations



## Computer Vision

- A branch of Artificial Intelligence
- Builds an AI that replicates human perception based analysis
- Is being used in numerous applications: for example driverless cars
- Very useful for capturing urban features and characteristics at scale
- Fast and Scalable – can scan all of urban Australia (60 million images, more than 20,000 square kilometers, in under 24 hours)

### Recreating and analyzing spatial patterns using imagery



Exploring the correlation of urban features with the spatial distribution of alcohol use



Detecting the spatial distribution of cycling infrastructure in Melbourne



# Urban infrastructure identification through computer vision and neural networks

Urban feature analysis from aerial remote sensing  
imagery using self-supervised and semi-supervised  
computer vision

Sachith Seneviratne<sup>a</sup>, Jasper S. Wijnands<sup>b</sup>, Kerry Nice<sup>a</sup>, Haifeng Zhao<sup>a</sup>,  
Branislava Godic<sup>a</sup>, Suzanne Mavoa<sup>a,f</sup>, Rajith Vidanaarachchi<sup>a,d</sup>, Mark  
Stevenson<sup>a,d,e</sup>, Leandro Garcia<sup>c</sup>, Ruth F. Hunter<sup>e</sup>, Jason Thompson<sup>a</sup>

## Self-Supervision, Remote Sensing and Abstraction: Representation Learning Across 3 Million Locations

Sachith Seneviratne<sup>\*</sup>, Kerry A. Nice<sup>\*</sup>, Jasper S. Wijnands<sup>\*</sup>, Mark Stevenson<sup>†\*</sup>, Jason Thompson<sup>\*</sup>

<sup>\*</sup>Transport, Health, and Urban Design Research Lab, Melbourne School of Design, University of Melbourne, Australia

<sup>†</sup>Melbourne School of Engineering, University of Melbourne, Australia

{sachith.seneviratne, kerry.nice, jasper.wijnands, mark.stevenson, jason.thompson}@unimelb.edu.au

Seneviratne, S., Wijnands, J. S., Nice, K., Zhao, H., Godic, B., Mavoa, S., Vidanaarachchi, R., Stevenson, M., Garcia, L., Hunter, R. F., and Thompson,

J. (2022b). [Urban feature analysis from aerial remote sensing imagery using self-supervised and semi-supervised computer vision.](#)

[arXiv](#)

Seneviratne, S., Nice, K. A., Wijnands, J. S., Stevenson, M., and Thompson, J. (2021). [Self-Supervision. Remote Sensing and Abstraction: Representation Learning Across 3 Million Locations.](#)

[2021 Digital Image Computing: Techniques and Applications \(DICTA\)](#)

# Models for urban infrastructure detection

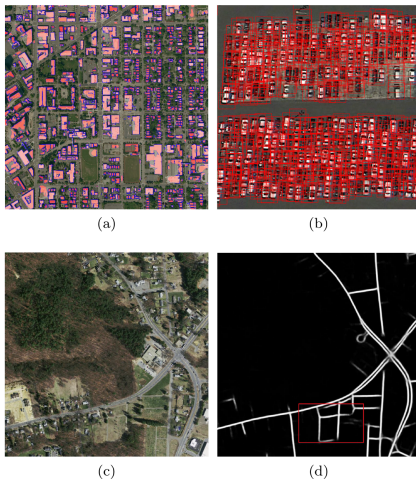


Figure 2: Specialised object detection models: (a) buildings [29], (b) vehicles [38], (c-d) roads

# Detection of cycling infrastructure in aerial imagery

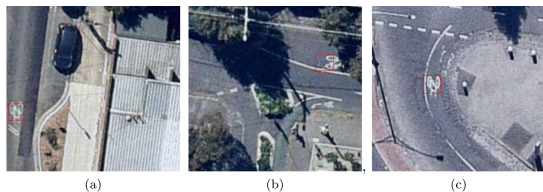


Figure 7: A red box is drawn around the region within the image most associated with the category under consideration (cycling symbols in this case) by the trained model

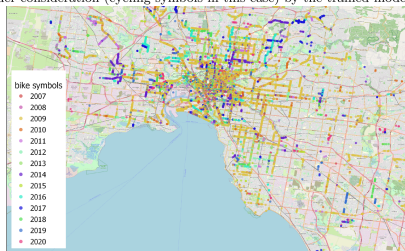


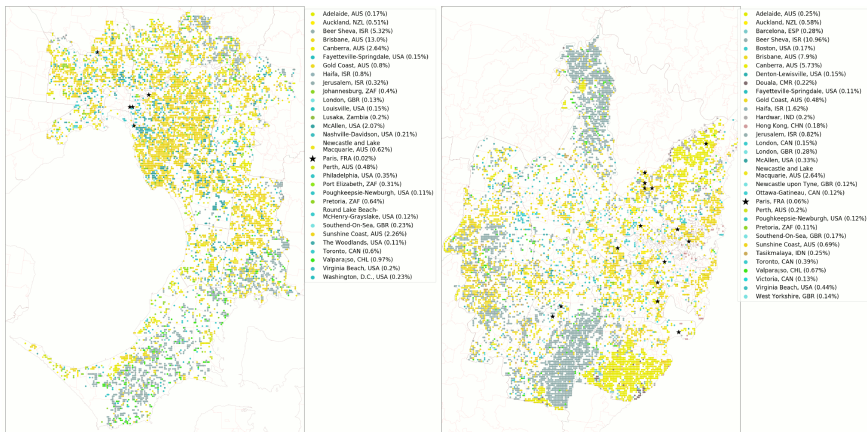
Figure 5: Generated GIS layer of cycling infrastructure over Melbourne

# Urban typologies through maps, satellite and street view images



**Figure 1.** Sample neural network training images for (a) GM (from Paris, France [51]), (b) GS (from Adelaide, Australia [51]), and (c) for GSV-BSV after mean shift pre-processing (from Sydney, Australia [52]).

# Inter and intra-city typologies



**Figure 3.** Predicted similar cities using the GM neural network (filtering out probabilities lower than 50%). Top predicted cities plotted (using Figure 2 colour scheme) for Melbourne evaluation locations (**left**) and Sydney (**right**). Predicted Paris locations marked with black stars.

# Urban typologies through maps and street view

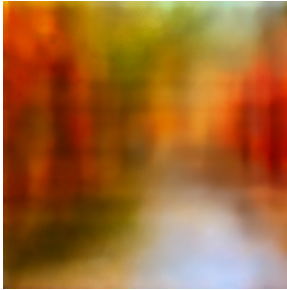


Figure 7. Gallery of “Paris-like” locations in Sydney using the GM neural network.



Figure 8. Gallery of “Paris-like” locations in Melbourne (A and B) and Sydney (C) using the GSV-BSV neural network.

Text-to-image: 'A street which is ideal for walking in'



# Text-to-image generation: Reimagining urban design

## Remembering and interpreting Parkville through artificial eyes

A/Prof Jason Thompson, Dr Sachith Seneviratne, Dr Kerry Nice, Dr Rajith Vidanaarachchi  
Transport, Health and Urban Design Research Laboratory – School of Design, The University of Melbourne

Could we conjure up an accurate vision of a time or a place if we had never been there? Could we then envisage such a place through another's eyes? What if we could delve into entire world's collective understanding of a place or perspective and extract an example of that representation? Would such a vision be robust, believable, or an any way meaningful? What effect might extracting such a scene have on our own understanding of empathy, creativity, imagination, possibility, or design practice?

To date, a limiting feature of artificial intelligence (AI) in applications ranging from those used in financial markets to health, object detection, and design has been their reliance on vast historical datasets from which they build perceptions and projections of the world before to relaying this information back to humans. This reliance has so far left scant room for apparent understanding, interpretation, or imagination of the sort truly recognised as 'intelligence' by humans. Such limitations, however, are perhaps drawing to a close with the advent of text-to-image generators(1). Text-to-image generators are opening the door for AI applications to reach human levels of conceptual creativity resembling considered integration, interpretation, and genuine imagination – with the capacity to even build memories of their own.

Contemporary text-to image generators are now enabling designers to combine and re-imagine features, objects, locations, times, scenes, and styles that can generate unique visions spanning anywhere from reality to impressionism to bizarre fiction. In doing so, they can lead us on an exploration of an alternate and at times fascinating potential world of how others might have seen or interpreted a place like Parkville.

In this presentation we demonstrate a re-imagining of Parkville and surrounds through capturing the collective hive-mind of the world-wide web using modern text-to image generators. We demonstrate an interpretation of Parkville under conditions where the features we desire and the lens we peer through are selected, combined, and balanced. We ask for consideration of what the model really understands, what the images convey, and leave open a discussion of the aesthetics, ethics, and meaning of artificially generated design to the audience. We show how artificial design resembling imagination, will be a feature of our future and perhaps eventually, even our history.



Figure 1. Four scenes from Parkville, Melbourne understood through a combination different styles and mediums, generated using both MidJourney and DALLE-2 text to image generators.



## DALLE-URBAN: Capturing the urban design expertise of large text to image transformers

Sachith Seneviratne\*, Damith Senanayake<sup>†</sup>, Sanka Rasnayaka<sup>‡</sup>, Rajith Vidanaarachchi\*<sup>†</sup>, Jason Thompson\*

\*Transport, Health, and Urban Design Research Lab, Melbourne School of Design, University of Melbourne, Australia

<sup>†</sup>Department of Mechanical Engineering, Faculty of Engineering and Information Technology, University of Melbourne, Australia

<sup>‡</sup>Department of Computer Science, National University of Singapore, Singapore

{sachith.seneviratne, damith.senanayaka, rajith.vidanaarachchi, jason.thompson}@unimelb.edu.au, <sup>‡</sup>sanka@nus.edu.sg

**Abstract**—Automatically converting text descriptions into images using transformer architectures has recently received considerable attention. Such advances have implications for many applied design disciplines across fashion, art, architecture, urban planning, landscape design and the future tools available to such disciplines. However, a detailed analysis capturing the capabilities of such models, specifically with a focus on the built environment, has not been performed to date. In this work, we investigate the capabilities and biases of such text-to-image methods as it applies to the built environment in detail. We use a systematic grammar to generate queries related to the built environment and evaluate resulting generated images. We generate over 12000 different images from 2 text to image models and find that text to image transformers are robust at generating realistic images across different domains for this use-case. Generated imagery can be found at the github: <https://github.com/sachith500/DALLEURBAN>

**Index Terms**—image generation, computer vision, deep learning, natural language processing, built environment, urban planning, urban design, dataset

### I. INTRODUCTION

The recent advent of deep learning methods that use large text-image datasets, combined using contrastive learning and transformer architectures has led to the possibility of generating realistic images of natural and artificial scenes, such as the image in Fig. 1. These methods have the potential

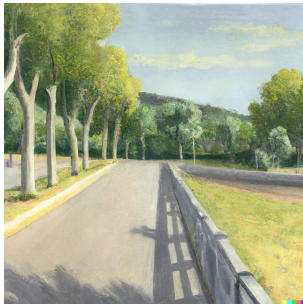


Fig. 1: DALL-E 2 generated image for the prompt: 'A painting of a public open space which is ideal for walking'

# Bias and capabilities of text-to-image generation related to built environments

TABLE V: In-painting as a means of validating the findings from text conditioned generation.

Healthy (a) → Unhealthy (b-d) (trees inpainted)



(a)



(b)



(c)

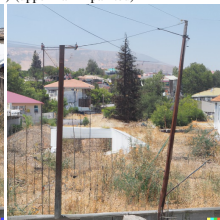


(d)

Unsafe (e) → Safe (f-h) (upper half inpainted)



(e)



(f)



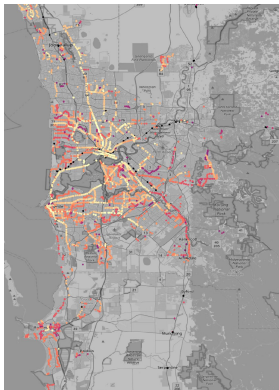
(g)



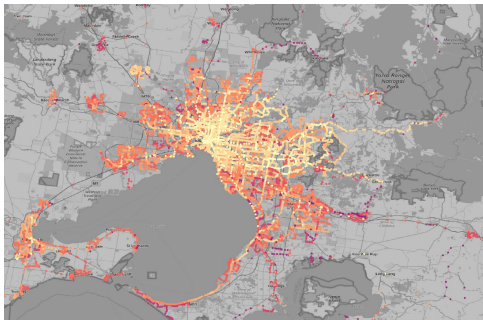
(h)

# Computer vision applications to derive a spatial index of access to social services

## Public transport availability and frequency



Services per week (lighter colors=higher frequency) in Perth and Melbourne.



# Computer vision applications to derive a spatial index of access to social services

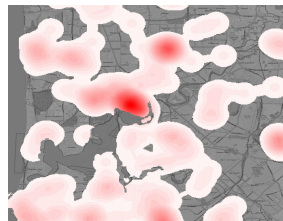
## Spatial access index creation

POI Classifier Tag	Categories of Amenities						
	Mobility	Active Living	Entertainment	Food	Community	Education	Health and Wellbeing
taxi_stop	X						
bus_station	X						
bus_stop	X						
public_transport	X						
bicycle_parking		X					
gym		X					
fitness_centre		X					
sports_centre		X					
park		X					
pitch		X					
playground		X					
swimming_pool		X					
garden		X					
pub_ground		X					
sports_centre		X					
ice_rink		X					
dog_park		X					
nature_reserve		X					
fitness_centre		X					
marina		X					
recreation_ground		X					
fitness_station		X					
skate_park		X					
pub			X				
bar			X				
theatre			X				
cinema			X				
nightclub			X				
gymnasium			X				
restaurant				X			
cafe				X			
food_court				X			
marketplace				X			
community_centre				X			
library					X		
social_facility					X		
social_centre					X		
townhall					X		
school						X	
childcare						X	
child_care						X	
kindergarten						X	
university						X	
college							X
pharmacy							X
dentist							X
clinic							X
hospital							X
doctors							X

Aggregate services and features into classes using different weights to create indexes

Use different mixes for different outcomes (i.e. young families, elderly, etc.)

Category	Weight
Mobility	0.2
Active Living	0.1
Nightlife	0.1
Food Choices	0.15
Community Space	0.1
Education	0.15
Health and Wellbeing	0.2

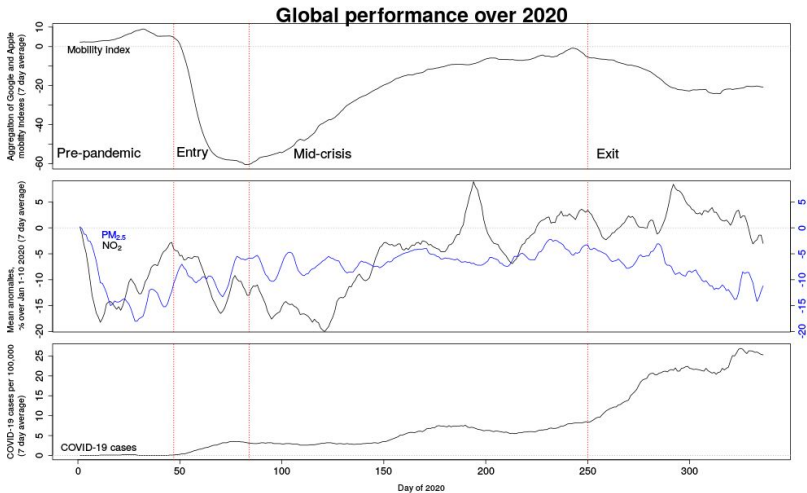


Supermarkets within 2000m



Doctors within 2000m

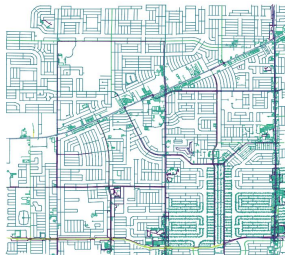
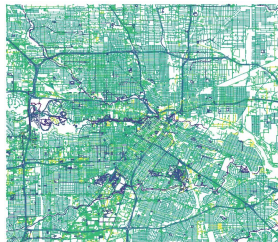
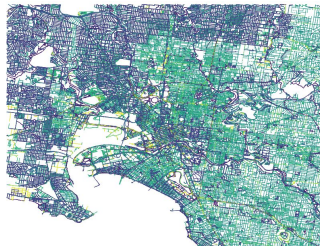
# COVID-19 over 2020: Global cities and responses constrained by urban design



# City network graphs to predict pollution outcomes

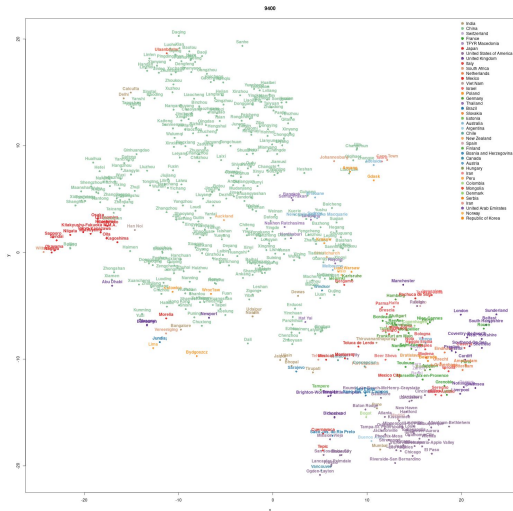
Clustering of cities using network graphs: transport nodes/edges, road types (primary, secondary, residential, etc) to predict 2020 pollution levels ( $\text{NO}_2$ ,  $\text{PM}_{2.5}$ )  
Melbourne (right), Houston TX, Cape Coral FL and Sanhe China (bottom row)

As well as other city characteristics (COVID cases, COVID stringency, etc across 2020).

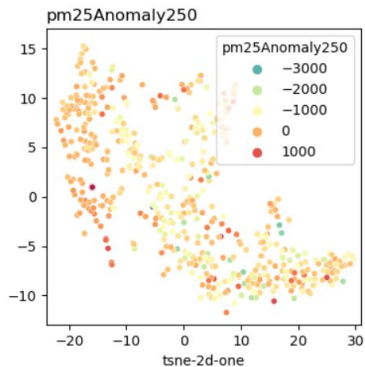
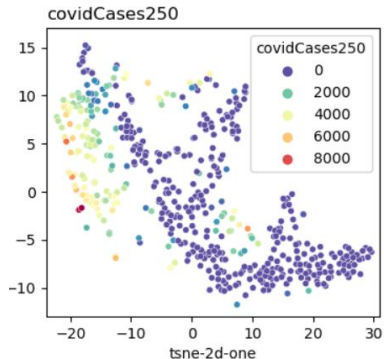


# Global response to COVID-19 over 2020

Clusters of cities by criteria of network graphs, stringencies during COVID and pollution outcomes.



## Health-related outcomes at day 250





## Transport, Health, and Urban Design Research Lab, Melbourne School of Design, University of Melbourne

<https://thud.msd.unimelb.edu.au/>

Dr. Kerry Nice (Software engineer, Urban climate modelling,  
<https://mothlight.github.io/> )

A/Prof. Jason Thompson (Psychology)

Prof. Mark Stevenson (Epidemiology)

Dr. Haifeng Zhao (Geoinformatics)

Dr. Sachith Seneviratne (Computer science)

Dr. Rajith Vidanaarachchi (Computer science)

Dr. Jasper Wijnands (Mathematics)



# References

- Hunter, R. F., García, L., Stevenson, M., Nice, K., Wijnands, J. S., Kee, F., Ellis, G., Anderson, N., Seneviratne, S., Moeinaddini, M., Godic, B., Akaraci, S., and Thompson, J. (2023). Computer vision, causal inference and public health modelling approaches to generate evidence on the impacts of urban planning in non-communicable disease and health inequalities in UK and Australian cities: A proposed collaborative approach. [medRxiv](#).
- Lipson, M. J., Nazarian, N., Hart, M. A., Nice, K. A., and Conroy, B. (2022). A Transformation in City-Descriptive Input Data for Urban Climate Models. [Frontiers in Environmental Science](#), 10:866398.
- Nice, K. A., Nazarian, N., Lipson, M. J., Hart, M. A., Seneviratne, S., Thompson, J., Naserikia, M., Godic, B., and Stevenson, M. (2022). Isolating the impacts of urban form and fabric from geography on urban heat and human thermal comfort. [Building and Environment](#).
- Nice, K. A., Thompson, J., Wijnands, J. S., Aschwanden, G. D. P. A., and Stevenson, M. (2020a). The "Paris-end" of Town? Deriving Urban Typologies Using Three Imagery Types. [Urban Science](#), 4:27.
- Nice, K. A., Wijnands, J. S., Middel, A., Wang, J., Qiu, Y., Zhao, N., Thompson, J., Aschwanden, G. D. P. A., Zhao, H., and Stevenson, M. (2020b). Sky pixel detection in outdoor imagery using an adaptive algorithm and machine learning. [Urban Climate](#), 31:100572.
- Seneviratne, S., Nice, K. A., Wijnands, J. S., Stevenson, M., and Thompson, J. (2021). Self-Supervision. Remote Sensing and Abstraction: Representation Learning Across 3 Million Locations. [2021 Digital Image Computing: Techniques and Applications \(DICTA\)](#).
- Seneviratne, S., Senanayake, D., Rasanayaka, S., Vidanaarachchi, R., and Thompson, J. (2022a). DALLE-URBAN: Capturing the urban design expertise of large text to image transformers. In [DICTA 2022](#).
- Seneviratne, S., Wijnands, J. S., Nice, K., Zhao, H., Godic, B., Mavoja, S., Vidanaarachchi, R., Stevenson, M., García, L., Hunter, R. F., and Thompson, J. (2022b). Urban feature analysis from aerial remote sensing imagery using self-supervised and semi-supervised computer vision. [arXiv](#).
- Stevenson, M., Thompson, J., Sá, T. H. D., Ewing, R., Mohan, D., McClure, R., Roberts, I., Tiwari, G., Giles-Corti, B., Sun, X., Wallace, M., and Woodcock, J. (2016). Land use, transport, and population health: estimating the health benefits of compact cities. [The Lancet](#), 6736(16):1–11.
- Thompson, J., McClure, R., Blakely, T., Wilson, N., Baker, M. G., Wijnands, J. S., de Sa, T. H., Nice, K., Cruz, C., and Stevenson, M. (2022a). Modelling SARS-CoV-2 disease progression in Australia and New Zealand: an account of an agent-based approach to support public health decision-making. [Australian and New Zealand Journal of Public Health](#).
- Thompson, J., Seneviratne, S., Nice, K., and Vidanaarachchi, R. (2022b). Remembering and interpreting Parkville through artificial eyes. In [Australian Centre for Architectural History, Urban and Cultural Heritage \(ACAHUH\) Inaugural Symposium : Park Life 2022, 04 November 2022](#).
- Thompson, J., Stevenson, M., Wijnands, J. S., Nice, K. A., Aschwanden, G. D. P. A., Silver, J., Nieuwenhuijsen, M., Rayner, P., Schofield, R., Hariharan, R., and Morrison, C. N. (2020). A global analysis of urban design types and road transport injury: an image processing study. [Lancet Planetary Health](#), 4:32–42.
- Wijnands, J. S., Nice, K. A., Seneviratne, S., Thompson, J., and Stevenson, M. (2022). The impact of the COVID-19 pandemic on air pollution: A global assessment using machine learning techniques. [Atmospheric Pollution Research](#), 13(6):101438.
- Wijnands, J. S., Nice, K. A., Thompson, J., Zhao, H., and Stevenson, M. (2019a). Streetscape augmentation using generative adversarial networks: insights related to health and wellbeing. [Sustainable Cities and Society](#).
- Wijnands, J. S., Scully, K., Zhao, H., Nice, K. A., Guo, J., Stevenson, M., and Thompson, J. (2020). Identifying safe intersection design through unsupervised feature extraction from satellite imagery. [Computer-Aided Civil and Infrastructure Engineering](#).
- Wijnands, J. S., Thompson, J., Nice, K. A., Aschwanden, G. D. P. A., and Stevenson, M. (2019b). Real-time monitoring of driver drowsiness on mobile platforms using 3D neural networks. [Neural Computing and Applications](#).