Streetscape augmentation using generative adversarial networks: Insights related to health and wellbeing

Jasper S. Wijnands[a,*], Kerry A. Nice[a], Jason Thompson[a], Haifeng Zhao[a], Mark Stevenson[a,b,c]

[a] *Transport, Health and Urban Design, Melbourne School of Design, The University of Melbourne, Parkville, VIC 3010, Australia*
[b] *Melbourne School of Engineering, The University of Melbourne, Parkville, VIC 3010, Australia*
[c] *Melbourne School of Population and Global Health, The University of Melbourne, Parkville, VIC 3010, Australia*

ABSTRACT

Deep learning using neural networks has provided advances in image style transfer, merging the content of one image (e.g., a photo) with the style of another (e.g., a painting). Our research shows this concept can be extended to analyse the design of streetscapes in relation to health and wellbeing outcomes. An Australian population health survey ($n = 34,000$) was used to identify the spatial distribution of health and wellbeing outcomes, including general health and social capital. For each outcome, the most and least desirable locations formed two domains. Streetscape design was sampled using around 80,000 Google Street View images per domain. Generative adversarial networks translated these images from one domain to the other, preserving the main structure of the input image, but transforming the 'style' from locations where self-reported health was bad to locations where it was good. These translations indicate that areas in Melbourne with good general health are characterised by sufficient green space and compactness of the urban environment, whilst streetscape imagery related to high social capital contained more and wider footpaths, fewer fences and more grass. Beyond identifying relationships, the method is a first step towards computer-generated design interventions that have the potential to improve population health and wellbeing.

1. Introduction

1.1. Streetscape design

Given the recent mass movement of people to cities, a trend expected to continue into the future (International Organization for Migration, 2015), neighbourhood planning is becoming increasingly important. Historically, as a response to problems related to overcrowding such as outbreaks of diseases following the industrial revolution, improving the health of city residents has been one of the main drivers of urban planning (Sharifi, 2016). For example, in the early 20th century the Garden City movement (Howard, 1902) aimed to address these issues by combining the best elements of city and countryside living in a single neighbourhood design. This was followed by the Neighbourhood Unit concept (Perry, 1929), attempting to address social problems by developing self-contained neighbourhoods that increase pedestrian safety and provide a strong sense of community. However, according to Mehaffy, Porta, and Romice (2015), the Neighbourhood Unit did not improve social interaction, health conditions or walking behaviour. Hence, although utopian visions aimed to

improve health and wellbeing outcomes of city residents, not all were successful. Modernism concepts, such as Le Corbusier's Radiant City (Fishman, 1982; Le Corbusier, 1933), introduced superblocks and high-rise buildings. In contrast, the Broadacre City (Wright, 1932) reserved one-acre plots per family for living purposes. Single-use zoning had some advantages, such as ease of implementation. However, suburban development led to increased car dependency and, in reality, had adverse health effects including non-communicable disease. Later, neo-traditional movements focussed on walkability, public transport, compact form and medium-high density to address issues caused by suburbanisation (Furuseth, 1997; Rutheiser, 1997). Eco-urbanism aimed for sustainable development, for example, using renewable energy technologies to reduce greenhouse gas emissions and protecting natural environments from new developments by revitalising existing urban areas instead (Joss, Cowley, & Tomozeiu, 2013; Tsolakis & Anthopoulos, 2015). Several design concepts have been linked to sustainability, including compactness, sustainable transport, density, mixed land use, diversity, passive solar design and greening (Jabareen, 2006). The neo-traditional development, eco-city and compact city (Dantzig & Saaty, 1973) all incorporate some of these sustainable urban

---

form concepts. These developments indicate that the historically strong link between urban form and health has been re-emerging.

Various recent studies have linked the urban form to health and wellbeing outcomes (de Vries, van Dillen, Groenewegen, & Spreeuwenberg, 2013; Sallis et al., 2015). For example, hard urban surfaces and reduced shading can lead to heat island effects, with associated health risks such as heat exhaustion and heat stroke, especially impacting the elderly (Carmona, Tiesdell, Heath, & Oc, 2010; Smith & Levermore, 2008). Urban green space has a cooling effect that could mitigate some of these issues (Bowler, Buyung-Ali, Knight, & Pullin, 2010). The level and variability of streetscape greenness has further health impacts, e.g., on perceived general health (Maas, Verheij, Groenewegen, de Vries, & Spreeuwenberg, 2006) and the prevalence of cardiovascular disease (Pereira et al., 2012). Survey-based approaches have identified that living in a green environment has positive associations with self-reported health, including the number of symptoms experienced in the last two weeks (de Vries, Verheij, Groenewegen, & Spreeuwenberg, 2003). Direct physical connections have been found as well, for example, Ward Thompson et al. (2012) showed that a lack of green space in deprived communities is associated with increased stress levels as measured by salivary cortisol.

Urban form also influences the frequency of active transport such as walking and cycling (Frank & Engelke, 2001; Hensley, Mateo-Babiano, & Minnery, 2014), providing associated health benefits (Turrell, Haynes, Wilson, & Giles-Corti, 2013). A higher perceived confinement of space and sense of intimacy (as measured by visual enclosure) is related to increased pedestrian activity (Yin & Wang, 2016). Specific streetscape features such as the proportion of windows on the street, the proportion of active street frontage and the amount of street furniture also significantly impact pedestrian activity (Ewing, Hajrasouliha, Neckerman, Purciel-Hill, & Greene, 2016). Further, pavement condition is considered very important by pedestrians and poor conditions lead to strong dissatisfaction among elderly pedestrians, which can affect physical activity levels (Stradling, Anable, & Carreno, 2007).

The built environment also affects crime rates and perceived safety (Foster & Giles-Corti, 2008). For example, Salesses, Schechtner, and Hidalgo (2013) found a significant association between violent crime and the perceptions of streetscapes related to safety and social class. Interventions that improve the appearance of abandoned buildings (e.g., by installing working doors and windows) and vacant lots (e.g., removing rubbish, planting grass and trees, installing low fences) can reduce the prevalence of firearm violence (Branas et al., 2016). The effect of greening vacant land also has mental health implications. Through a cluster randomised trial, South, Hohl, Kondo, MacDonald, and Branas (2018) found that greening significantly reduced feelings of depression and worthlessness of residents living nearby. Cohen, Inagami, and Finch (2008) found that parks are positively associated with the perception of mutual trust and willingness of people to help each other. Further, the architectural features present in urban streetscapes have been found to affect wellbeing outcomes (Spokane et al., 2007). A literature review by Mair, Diez Roux, and Galea (2008) identified that depression was more consistently associated with measures of the built environment than with socioeconomic deprivation, residential stability or race composition.

The studies above indicate the perception of urban form has various health and wellbeing impacts. Lynch (1960) investigated the human perception of urban form and concluded humans map their surroundings using elements such as paths, edges, districts, nodes and landmarks. Further, Lozano (1990) defined urban form as the aggregation of more or less repetitive elements, such as block size and form, street design, street patterns and the layout of parks. Big data sources such as Google Street View (GSV) imagery (Anguelov et al., 2010) now allow for in-depth analyses of these urban features. For example, large numbers of geo-tagged photos have been used to detect patterns of urban usage and public perception of functional and social attributes (Liu, Zhou, Zhao, & Ryan, 2016; Zhou, Liu, Oliva, & Torralba, 2014).

Doersch, Singh, Gupta, Sivic, and Efros (2012) used geo-localised street level images to discover unique visual features in a city, such as balconies, windows with railings and special Parisian lampposts in Paris, France. Place Pulse, a database of urban imagery using crowd-sourced classifications, quantifies perceptions of urban areas, including safety, beauty, and liveliness (Dubey, Naik, Parikh, Raskar, & Hidalgo, 2016; Naik, Philipoom, Raskar, & Hidalgo, 2014). Various studies have also attempted to quantify the level of green space visible at the street level. For example, Li et al. (2015) used colour information in GSV images to estimate and map urban green space. Further, Seiferling, Naik, Ratti, and Proulx (2017) focussed on identifying tree cover from GSV images using computer vision techniques.

Although the research findings described throughout this introduction are important, the methods employed typically focus on a selection of urban factors or sometimes just one factor in isolation (e.g., greenness). As these factors are pre-supposed to be associated with outcomes, they are potentially subject to a degree of observer bias based on historical, theoretical, or disciplinary background. Further, when multiple factors are included, there is limited understanding of how these interact to improve health or how suggested changes to urban form might translate in the real world. For example, while traditional linear approaches such as regression models might include multiple observed factors that recommend a '30% improvement in greenness', '25% increase in density', and a '15% increase in diversity', it is left to the imagination of residents, policy-makers and planners to understand how these changes might actually appear when translated into the urban landscape. Given the mixed success, historically, of utopian planning visions on city residents' health, the analysis of empirical streetscape data using an objective, unsupervised approach will identify elements of neighbourhood planning movements that can be linked to actual health improvements of residents.

1.2. Style transfer

Compared to the traditional approaches described above, the current study does not require an a priori selection of features potentially subject to observer bias. Our approach is based on the concept of generative adversarial networks (GANs), first described in Goodfellow et al. (2014). GANs are based on game theory and consist of multiple competing models, namely generator and discriminator neural networks. The role of the generator is to generate images in the new style, while the discriminator assesses whether the generated image looks realistic. Using different formulations of the internal loss function and neural network structures, GANs have been adapted for various image translation tasks; for example, increasing image resolution to recover finer details (i.e., super-resolution) (Ledig et al., 2017; Sønderby, Caballero, Theis, Shi, & Huszár, 2017) or filling in missing regions of an image (i.e., semantic inpainting) (Pathak, Krähenbühl, Donahue, Darrell, & Efros, 2016; Yeh et al., 2017). Another application of image-to-image translation is style transfer, which entails merging the style of a collection of images and the content of another image. Examples of style transfer include translating (i) photos to paintings, (ii) black and white photos to colour, (iii) summer photos to winter and (iv) daytime photos to night (e.g., Gatys, Ecker, & Bethge, 2016; Isola, Zhu, Zhou, & Efros, 2017; Zhu, Park, Isola, & Efros, 2017).

Importantly, these techniques can also be applied to images representing the design of the urban environment. As health and wellbeing outcomes have been shown to be related to the human perception of urban form (see Section 1.1), the image style could be defined as streetscapes with either a positive or negative human perception. Salesses et al. (2013) have shown the large variety in human perceptions of streetscapes captured in GSV images. By changing the style of the streetscape towards a more positive perception, it might be possible to also improve health or wellbeing outcomes. Note that not all variation in GSV images may be attributed to human perception or health and wellbeing outcomes. While for the style transfer domains described

above (e.g., black and white to colour) the style is uniquely defined, the exact streetscape style that influences human perception and health outcomes still has to be discovered. Therefore, research is required that investigates if a GAN is able to successfully capture these styles.

Using GANs it may be possible to design by example rather than designing each urban space individually, based on large streetscape datasets. By sampling design of successful areas using outcomes such as liveability, non-communicable disease and mental health, unsupervised image-to-image translation may be able to apply key characteristics of these areas to any other location. This study provides an investigation along these lines, focussing on the differences in streetscape imagery from areas with good and bad perceived health and wellbeing outcomes. The proposed research has two objectives: (i) to present the application of a new technique for streetscape analysis based on big data and artificial intelligence; and (ii) to present findings of relationships between health outcomes and specific design elements discovered using this method.

2. Methods

This study used GANs to augment the design of streetscapes captured in GSV images, as GSV provides an impression of streetscape design from an observer's point of view. The locations at which to obtain these GSV images were determined using survey-based health and wellbeing data.

2.1. Health and wellbeing

The data used to identify the health, wellbeing and demographic characteristics of areas of the Melbourne metropolitan region were derived from a Victorian Population Health Survey (Department of Health & Human Services, 2017). This health survey is a population-weighted computer-aided telephone interview survey that collected information from 34,000 residents across the State of Victoria relating to self-reported health outcomes and health behaviours. For this research, health outcomes and behaviours relating to perceptions of general health, social capital and life satisfaction across metropolitan Melbourne were included. Specifically, individuals were asked to rate their perceptions of general health on a five-point Likert scale ranging from 1 (excellent) to 5 (poor). Social capital was measured by asking individuals to recall the number of people they had spoken to on the day before. Overall life satisfaction was estimated by asking individuals how satisfied they were with life overall, with ratings from 1 (very satisfied) to 4 (very dissatisfied). This was supplemented with population and housing density data for validation purposes.

2.2. Google Street View

A GSV dataset was created by sampling the Greater Melbourne area, Australia. Locations were determined using the nodes of vector lines in a street network dataset (PSMA Australia Ltd., 2017) to exclude most indoor imagery. For each selected location, four GSV images were retrieved at headings of 0, 90, 180 and 270 degrees and the Google logo was removed for deep learning purposes. A sample of 100,000 images was inspected for consistency and the complete dataset was post-processed by removing remaining indoor images based on file size, resulting in 4.5 million GSV images at a 256×256 pixels resolution. This dataset and the underlying methodology is explained in more detail in the linked *Data in Brief* paper (Nice and Wijnands, 2019).

For each outcome variable of interest, the 10% most and least desirable locations were selected with respect to self-reported health outcomes (as defined above) and were grouped into two sets (i.e., domains). The resulting locations were used to select a corresponding image from the GSV dataset. If no GSV image was available within 50 m of the requested location, the location was excluded, leading to approximately 160,000 images per outcome variable (i.e., 80,000 per domain). These images were used as input to the GAN.
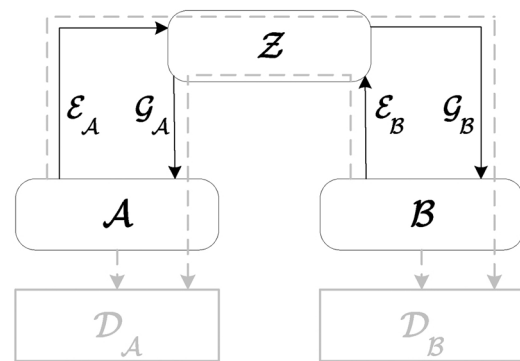

Fig. 1. Illustration of model configuration.

2.3. Generative adversarial networks

Since no matched pairs of images exist, supervised learning is not feasible. However, GANs have proven to be successful for style transfer using unsupervised learning (e.g., Zhu, Park, et al., 2017). Previous studies have considered 'style' as generating black and white images in colour, or as a painting (see Section 1.2). In contrast, this study defines style as the health or wellbeing outcome so images can be translated, for example, from a bad to a good general health style (as defined in the population health survey). The main structure and layout of the GSV image can be considered the 'content', such that any differences between the original and generated image could eventually be considered amendments that can be incorporated into the current urban environment (in future research).

Our approach is based on unsupervised image-to-image translation (Liu, Breuel, & Kautz, 2017) and consists of two variational auto-encoders (VAEs) (Kingma & Welling, 2013) plus two discriminator models. The VAEs create a condensed representation of an input image in a latent space ($\mathcal{Z}$), which is assumed to be shared between the two domains ($\mathcal{A}$ and $\mathcal{B}$). The model configuration is illustrated in Fig. 1. Encoding functions $\mathcal{E}_{\mathcal{A}}$ and $\mathcal{E}_{\mathcal{B}}$ extract features from images of domain $\mathcal{A}$ and $\mathcal{B}$, respectively. Generator functions $\mathcal{G}_{\mathcal{A}}$ and $\mathcal{G}_{\mathcal{B}}$ decode the condensed representation from $\mathcal{Z}$ to the respective domain. Hence, the content is the information stored in $\mathcal{Z}$, while the style is embedded in generator functions $\mathcal{G}_{\mathcal{A}}$ and $\mathcal{G}_{\mathcal{B}}$. For example, elements that occur in GSV images of both domains (e.g., people) will likely be stored in $\mathcal{Z}$ without large differences between $\mathcal{G}_{\mathcal{A}}$ and $\mathcal{G}_{\mathcal{B}}$ (although this is not directly enforced), while elements unique to one of the image sets (e.g., green space) will only be captured in the generator function for that specific domain.

In this model configuration, images can be encoded in the shared latent space and decoded in either the original domain (i.e., a perfect model will reconstruct the same input image) or the alternative domain. Ideally, images would also be accurately regenerated when processed from $\mathcal{A}$ to $\mathcal{Z}$, to $\mathcal{B}$, to $\mathcal{Z}$, to $\mathcal{A}$ and, similarly, $\mathcal{B} \rightarrow \mathcal{Z} \rightarrow \mathcal{A} \rightarrow \mathcal{Z} \rightarrow \mathcal{B}$ (cycle consistency). The discriminator models ($\mathcal{D}_{\mathcal{A}}$ and $\mathcal{D}_{\mathcal{B}}$) assess the quality of images translated to the alternative domains. The associated loss function combines all discriminator, reconstruction and cycle consistency losses using hyperparameters λ_0, (λ_1, λ_2) and (λ_3, λ_4), respectively (see Liu et al., 2017). These hyperparameters are used to control which of the loss components are emphasised during model training.

Various experiments with different hyperparameter values were performed to achieve a balanced translation. For example, when prioritising generator losses (either reconstruction or cycle consistency) translated images were very similar to the original, which does not allow for the identification of relationships between streetscape design and health outcomes. In contrast, when style transfer is too strong, the resulting image may appear abstract or the changes cannot be considered 'amendments' to the urban environment, but a complete

redesign. Compared to the default values, the selected hyperparameters $((\lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = (50, 0.1, 100, 0.1, 100))$ place more emphasis on the discriminator loss, prioritising the generation of more realistic images in the style domain over perfect reconstruction and cycle consistency.

Pre-processing steps included the reduction of image resolution to 128×128 pixels to speed up computation. The use of higher-resolution images with GANs is possible (Curto, Zarza, De La Torre, King, & Lyu, 2017), but was outside the scope of this study. GANs were then calibrated using backpropagation for 1.2 million iterations (approximately 60 h of training on an NVIDIA GTX 1080 GPU), with alternating weight updates to the generator and discriminator networks.

2.4. Analysis of generated images

The calibrated model was used to translate images from one domain to the other. Generated images were analysed in the following three ways: (i) a visual comparison using difference images; (ii) using statistical measures; and (iii) the computation of the average translation, providing insights into the neural network's inference patterns.

Visual comparison. To compare a generated image to its original, pixel-based difference images were constructed using ImageMagick (ImageMagick Studio LLC, 2018). The visual analysis of a large sample of translated images identified several themes per outcome variable. In Section 3, eight translated images are presented for most of the outcome variables, which are representative of these themes.

Statistical measures. Statistical analysis was used to estimate the strength of the association between investigated outcome variables and local urban form. Several hundred images for each combination of outcome (i.e., general health, social capital, life satisfaction) and domain (i.e., low or high) were randomly selected. Each image was translated to the opposite domain and a difference image was constructed. Since the GAN first creates a high-level representation of an image in the shared latent space, details such as the exact colour value and positioning are lost. Therefore, even if part of the image is unaffected by the new style, almost all pixel values of the image change during translation. In order to identify the extent to which an image was changed, small changes were discarded (ImageMagick *convert* function with parameters compose:difference, fuzz:5%, transparent:-black). The average proportion of non-white pixels gives an indication of the amount of change observed using the calibrated model.

The difference between the original and translated images was also investigated using standard statistical approaches; specifically, the mean squared error (MSE), peak signal-to-noise ratio (PSNR) (e.g., Bing, 2015) and structural similarity index (SSIM) (Wang, Bovik, Sheikh, & Simoncelli, 2004). These statistics compute the average similarity between the original and translated image. MSE and PSNR focus on pixel-based differences, while SSIM is geared towards the degradation of structural information. Specifically, SSIM combines structure, luminance and contrast measures for many small, corresponding patches in both images. The GAN translates the style of the GSV image and attempts to preserve its content at the same time. Therefore, it is expected that perceived changes in structural information will be limited by construction, while pixel-based differences could be more indicative of the amount of style transfer. For example, translation of a road surface from gravel to sealed road will lead to large pixel-based differences captured by MSE and PSNR, while the image structure could still be quite similar.

Average translation. Besides quantifying the amount of change occurring during the translation process, it is also possible to get a better understanding of the type of changes the GAN makes. Therefore, the average translation by the generator neural network was computed using all domain images. This can provide evidence on whether the identified themes hold overall. First, the average GSV image was computed pixel-by-pixel for corresponding pixels of all 80,000 images in a domain. Then, all these images were translated using the calibrated

GAN and the average translated image was computed. Differences between the average original and translated image give an insight into what the network prioritises when converting one style to another and exemplifies the inner workings of the generator neural network.

These three different methods are employed as the validity of results may be difficult to assess, since the theoretical foundation of associations between urban form and health and wellbeing outcomes is not captured by the proposed methodology (i.e., the method is based on large-scale empirical data). In addition, the face validity (e.g., consistency with expected performance) of the model is first investigated using domains that are readily understood: (i) density; and (ii) city/park translations.

3. Results

3.1. Visual comparison – density, city/park

Density translation. Fig. 2a shows the selected locations corresponding to the highest and lowest density in the Greater Melbourne area, with the highest density observed near Melbourne's central business district (CBD). After model training, the GAN reconstructs images accurately when encoding in the latent space and regenerating (results not presented). Fig. 3 shows the original GSV and translated images from low to high-density and vice versa. Specifically, the translation from low to high-density showed the following themes: (i) the creation of new buildings, generally in areas currently occupied by trees; and (ii) the conversion of grass and gravel surfaces to concrete (see Fig. 3a–d). The reverse translation of high to low-density can be characterised by (i) the removal of buildings, cars, street lights, road markings and road signs, which are generally replaced by grass, trees or bushes; (ii) the conversion of multi-level buildings to single-level; and (iii) the creation of larger open spaces or plains (see Fig. 3e–h).

Park and city. A more extreme transformation is the translation between Melbourne's inner-city and parks (i.e., Kings Domain and the Royal Botanic Gardens), as pictured in Fig. 2b. These domains have fewer images than the survey-based domains, namely about 6200 city and 13,800 park GSV images. Various themes were observed for the park to city translation, including (i) the reduction of open spaces by creating high-rise buildings; and (ii) the replacement of natural surfaces such as grass and gravel by concrete (see Fig. 4a–d). Other features, such as cars, trees and pedestrians, are maintained. The reverse translation also shows several themes, namely (i) the generation of additional green space; (ii) the conversion of sealed surfaces to gravel; and (iii) the elimination of vehicles, buildings and power lines, which are scarcely observed in park GSV images (see Fig. 4e–h).

Fig. 4h shows the translation of a large facade to multiple trees, although the translated image still contains remnants of the original building. Importantly, this indicates that large facades do not (or, rarely) occur in the park domain, leading to complete replacement.

The changes described above are in line with expectations when translating between low and high-density, and city and park domains, respectively. Translating from low to high-density leading to the addition of buildings is self-evident, as is building removal for the opposite translation. Similarly, the creation of natural surfaces in a park style and the addition of high-rise buildings in an inner-city style is clear. Visually the results have high face validity, providing confidence in the proposed method. Specifically, these experiments highlight that GANs can capture differences between selected image sets as distinct styles, even though streetscapes are heterogeneous and the image sets contain significant overlap (e.g., low-density areas may have some multi-level buildings) and not every element in the GSV image may be related to concepts as 'city' and 'park'. These results allow us to move forward and explore the transformation of health-related domains, where the captured styles are less obvious.
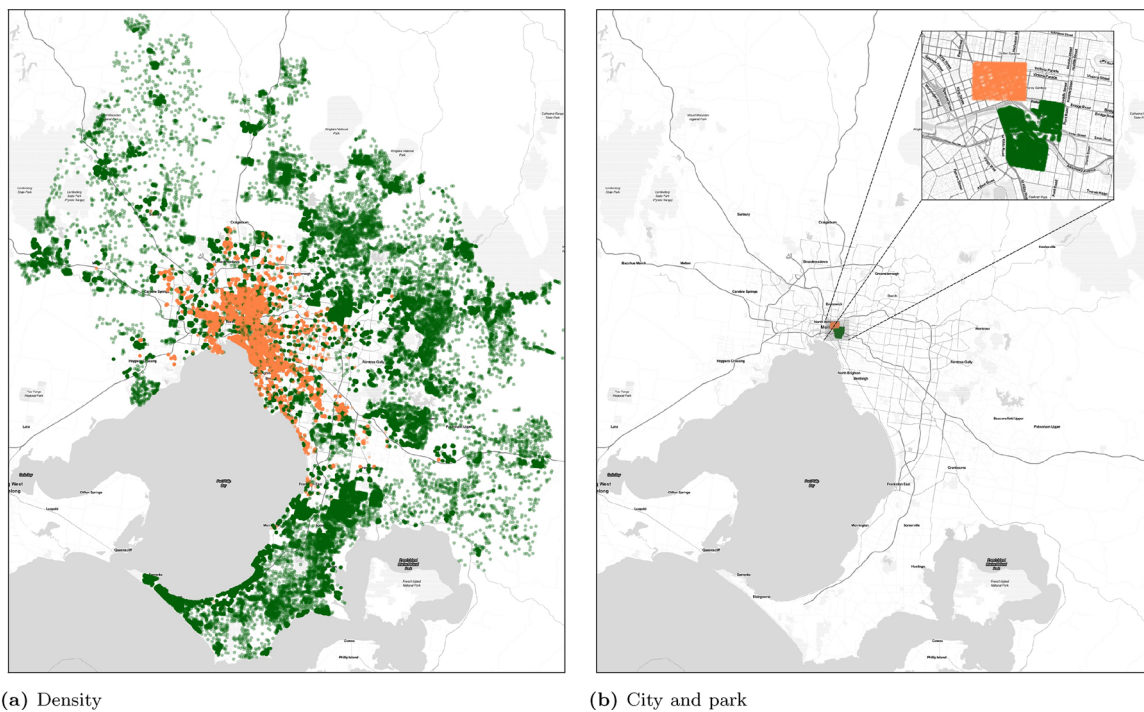
(a) Density

(b) City and park

Fig. 2. Selected domains: (a) locations with the highest (red/light grey) and lowest (green/dark grey) 10% of density, and (b) city (red/light grey) and park (green/dark grey) locations. Figures were generated using QGIS (QGIS Development Team, 2018). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
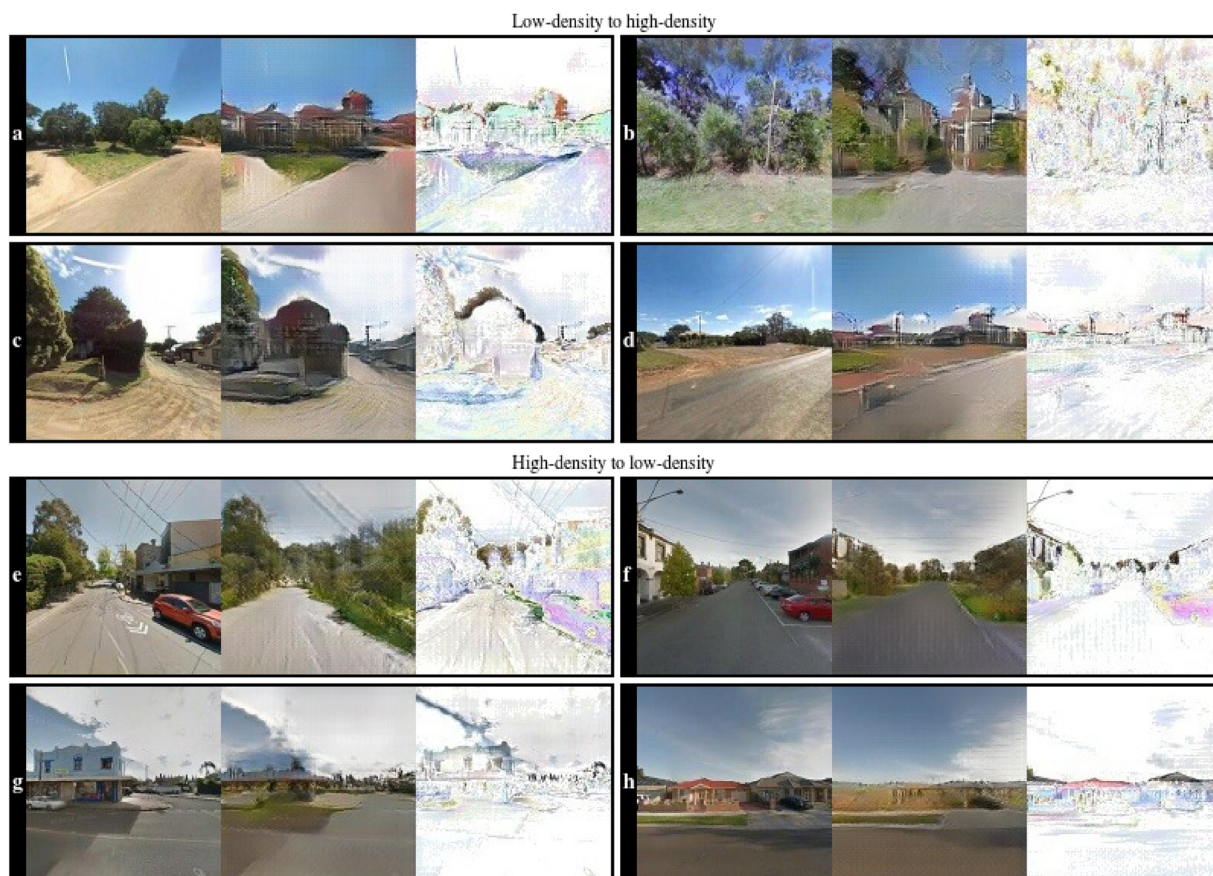

Fig. 3. Sample images showing translation of low-density GSV images to a high-density style (sub-figures a–d) and vice versa (e–h). Each sub-figure shows the original GSV image (left; source of map data: Google), the generated image (middle) and their pixel-by-pixel differences (right).

Fig. 4. As Fig. 3, but for the translation of park GSV images to a city style (sub-figures a–d) and vice versa (e–h).

3.2. Visual comparison – health and wellbeing

The selected locations for general health, social capital and life satisfaction are presented in Fig. 5. Some clustering of the domains is apparent, especially for general health and social capital. Self-reported general health is better in the CBD and coastal suburbs, while social capital is higher in the suburbs. For life satisfaction, the locations corresponding to the best responses are distributed over the Greater Melbourne area.

General health. The translation from a good to a bad general health style shows that buildings appearing close to the GSV camera are moved further away (see Fig. 6a–d). This indicates that compactness of the urban environment in Melbourne is mostly observed in areas with good general health. GSV images with nearby structures are generally not present in the bad general health domain, causing buildings to be moved and sky to appear during the translation. Fig. 6e and f exemplify the reduction in green space for the same translation, where trees become smaller and grass areas are reduced or removed. Natural surfaces, such as gravel footpaths in Fig. 6f, become sealed. The reverse translation (Fig. 6g and h) supports these findings, showing more grass and larger, denser trees.

Social capital. The translations related to social capital show that grass areas disappear, while trees become taller and more dense, when translating to the style based on low social capital GSV images (see Fig. 7a–d). The opposite pattern is apparent in the reverse translation, as exemplified in Fig. 7e. Interestingly, the identified themes for general health show a change in the amount of green space (i.e., grass, bushes and trees), while this is not the case here: *more* grass and *smaller* trees appear in GSV images related to high social capital. Wider and/or new footpaths (see Fig. 7e and f) and the removal of fences (see Fig. 7g and h) indicate that the selected image sets also differ for these features.

Life satisfaction. For life satisfaction no clear themes were observed, except a somewhat similar green space pattern as for social capital.

3.3. Statistical measures

Table 1 shows the amount of change (i.e., the proportion of non-white pixels in difference images) observed following translation. Further, Table 2 investigates the difference between the original and translated images using MSE, PSNR and SSIM. Note that an image-to-image translation showing many changes corresponds to a high MSE, low PSNR and low SSIM. Both Tables 1 and 2 show that the largest changes occur for the park to city conversion, with on average the highest proportion of changed pixels, highest MSE and lowest PSNR. These domains contain minimal overlap (i.e., most images from the city domain are distinctly different from the images in the park domain), which results in larger changes when translating an image from park to city and vice versa. Other domains, especially life satisfaction, show more overlap (also see the lower degree of clustering in Fig. 5c). For life satisfaction, the translated images are most similar to the original images with, on average, the lowest proportion of changed pixels, lowest MSE and highest PSNR. This is in line with the lack of themes identified for this variable during the visual analysis of the generated images (see Section 3.2). Therefore, it seems that life satisfaction is not as strongly related to local urban form as the other investigated outcome variables.

3.4. Average translation

Fig. 8 shows the average translations for general health and social capital. The changes in red, green and blue channels have been amplified to better visualise the spatial spread of changes in each channel.
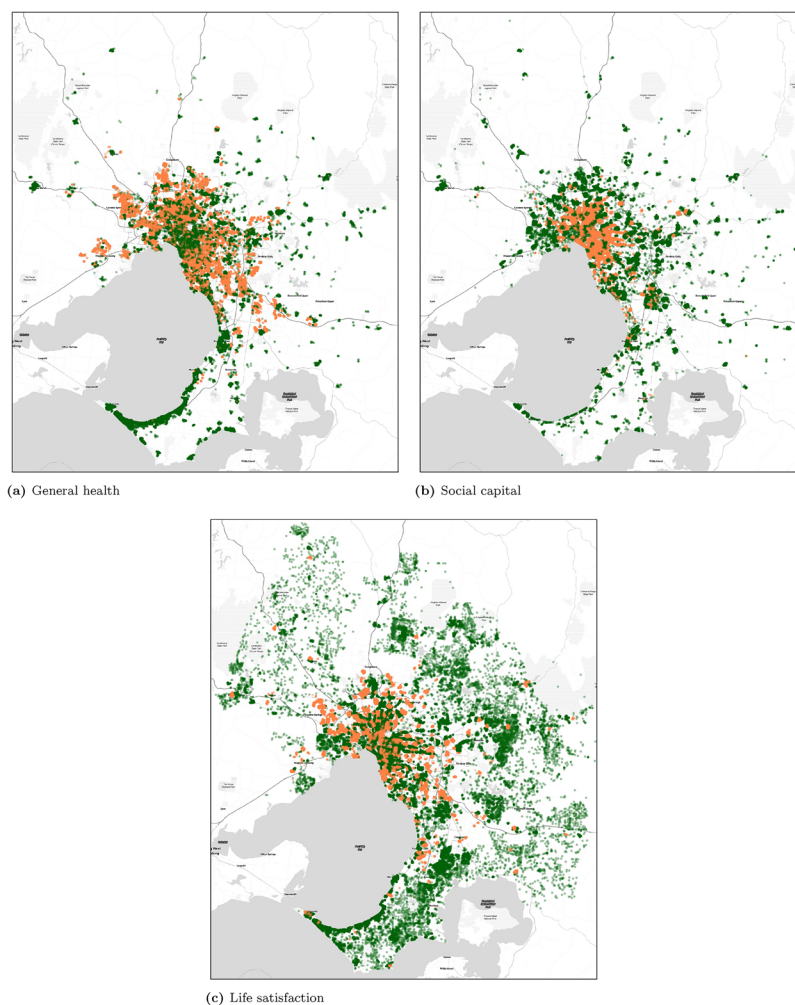
(a) General health

(b) Social capital

(c) Life satisfaction

Fig. 5. Selected domains: locations with the best 10% (green/dark grey) and worst 10% (red/light grey) of (a) general health, (b) social capital, and (c) life satisfaction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The RGB differences for the translation of images from bad general health locations to a good health style show more natural colours (i.e., yellow, green) below the horizon, supporting the findings in the visual analysis on the generation of additional grass and natural surfaces. The opposite translation shows more blue above the horizon, consistent with the visual analysis where buildings close to the GSV camera were partly replaced by sky and moved further away. The average translation to a high social capital style indicates changes mainly occur near the horizon, with more grass and more sky because of changes in vegetation (in line with the visual analysis). Finally, the reverse translation was linked to the replacement of grass areas by concrete, which is also exemplified by the blue colours below the horizon in Fig. 8.

Similar to the analysis above, the average pixel value and change in red, green and blue channels were calculated for all outcome variables and translations (see Table 3). The statistics show an increase in the green (and red) channels for translations towards styles related to low density, parks, good general health, high social capital and high life satisfaction. This confirms various findings of the visual analyses regarding increased green space (see Section 3.1 and 3.2).

4. Discussion and conclusion

This study presented an innovative approach to understanding variations in streetscape design through computer-generated visualisations. As described in the introduction (see Section 1.1), the ideal urban form has been a point of discussion for more than a century.

Utopian visions have been formulated as a response to issues of that time, leading to new developments or revitalisation of current urban areas. The impacts of these new conceptual frameworks are difficult to evaluate and are generally only visible after many years. The results presented in this paper are not derived from theory-driven urban planning concepts, but from a completely different perspective: the empirical exploration of unique features in large streetscape image datasets using artificial intelligence. When translating images to a new 'style', the method automatically focusses on features that differentiate the image sets of selected domains (i.e., unsupervised learning). These translations show that areas in Melbourne with good general health are characterised by sufficient green space, compactness of the urban environment and natural surfaces. Further, streetscape imagery representative of neighbourhoods where residents had high levels of social capital contained more and wider footpaths, fewer fences and more grass. The themes observed in the translated imagery, such as compactness and greening, are pointing towards sustainable design philosophies as the compact city and eco-city as viable ways to enhance the health of city residents.

To identify if GANs are capable of capturing health and wellbeing styles, the analysis was expanded by including statistical measures and the average translation, supporting the conclusions. Findings are also in line with previous research, providing confidence in the new method and the validity of its results. For example, the positive impact of green space on general health was also identified by various other studies (e.g., Groenewegen, van den Berg, de Vries, & Verheij, 2006; Maas

Fig. 6. As Fig. 3, but for the translation of GSV images at locations with good general health to a style related to bad general health (sub-figures a–f) and vice versa (g and h).

et al., 2006; Tzoulas et al., 2007). In addition, the positive health impact of compactness of the urban environment was quantified by Stevenson et al. (2016), finding compact city design yields health gains of 420–826 disability-adjusted life-years per 100,000 population. With respect to non-compact urban form, various studies indicate urban sprawl has negative effects on public health, caused by increases in air pollution, transport-related crashes and reduced physical activity (e.g., Frumkin, 2002; Giles-Corti, Ryan, & Foster, 2012). In contrast, the relationship between urban sprawl and social capital is less clear. Sprawl could support some types of social capital while negatively impacting others (Nguyen, 2010), indicating why compactness of the city was not observed in the social capital analyses. More green and other outdoor spaces, on the other hand, are associated with higher social capital. For example, Sullivan, Kuo, and Depooter (2004) found that suitable outdoor green space leads to a significant increase in the amount of social activity that takes place in these areas. Further, community gardens lead to increased social cohesion and social support (Kingsley & Townsend, 2006). Perceived greenness of the urban environment is correlated with recreational walking, social cohesion and local social interaction (Sugiyama, Leslie, Giles-Corti, & Owen, 2008). Further, Leyden (2003) found that walkable neighbourhoods lead to higher social capital. This is consistent with our findings relating social capital to outdoor space and infrastructure (i.e., grass and footpaths), although findings related to trees could be indicative of more remote areas.

This indicates a limitation of this study, as our method (as many other studies) does not provide causality of the identified relationships. For example, for both general health and social capital changes in green space were found, while Maas, van Dillen, Verheij, and Groenewegen (2009) indicate that loneliness and perceived shortage of social support could be a mediator for the relation between green space and health.

Outcomes such as general health are influenced by many factors; the design of streetscapes is not their only driver. For example, the urban form at a larger scale, including land use mix and access networks, is not captured in a streetscape image. In addition, many factors not related to design (e.g., genetics, education, diet) influence these outcomes. However, if there would be no relationship between the design of streetscapes and health and wellbeing outcomes, changes between the original and generated images would be minimal without any clear, outcome-dependent themes. As shown in this research, this is not the case. In addition to imagery, future studies could aim to explicitly incorporate these other factors (e.g., distance to public transport, amenities, neighbourhood walkability) in the modelling. Various studies stress the need to change policies and amend local urban environments as a method to improve health and safety, as the impact of local environment improvements may be larger than interventions targeted at individuals (e.g., Kondo, South, & Branas, 2015). Eventually, GAN-based methods could be capable of suggesting amendments to specific streetscapes that improve health and wellbeing outcomes, by generating targeted interventions that also retain the current infrastructure of the local environment.

Note that images in this study were obtained for Melbourne, Australia and applicability of the findings to other international cities has not been investigated. Validity of these findings beyond Melbourne could be investigated using a similar approach, but different input imagery and health outcome data. A further limitation is that not all images are of great quality and some contain artefacts; for example, the reconstruction of cars could be more accurate. Inaccuracies in these renderings do not influence the major themes identified through visual analysis in this research.

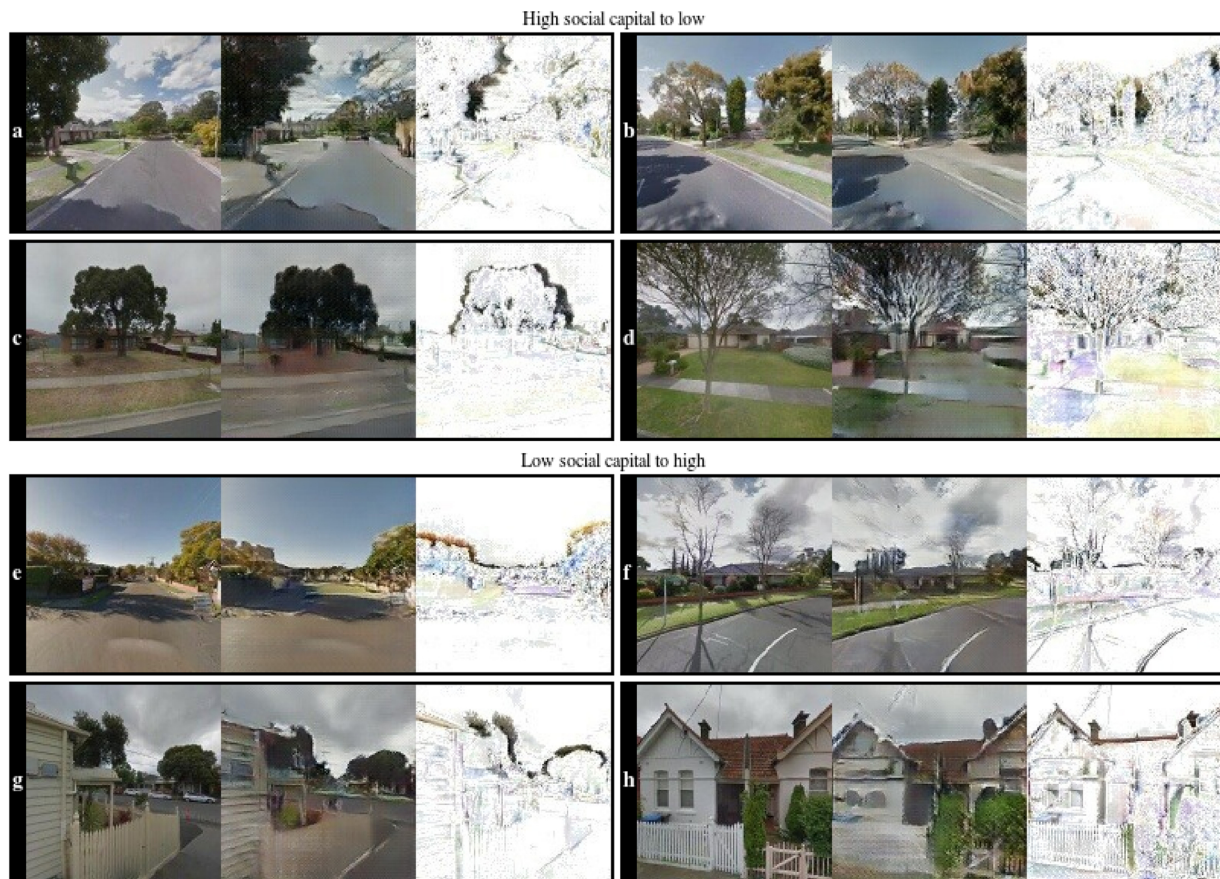We identify several additional avenues for future research. First,

Fig. 7. As Fig. 3, but for the translation from high to low social capital (sub-figures a–d) and vice versa (e–h).

Table 1

Average proportion of pixels that changes during translation.

	Low to high	High to low
Density	53.2%	57.8%
Park/city	86.1%	80.9%
General health	44.2%	54.2%
Social capital	42.2%	39.5%
Life satisfaction	35.5%	37.3%

GAN architectures based on multimodal unsupervised image-to-image translation (Huang, Liu, Belongie, & Kautz, 2018; Zhu et al., 2017b) can generate multiple realisations based on a single input image, as there are generally multiple valid solutions when translating to a different domain. For the theme analysis performed in our study, multimodal translation was not required, but it could be explored in future research. Second, the Place Pulse 2.0 database (Dubey et al., 2016) captured the perception of GSV images using crowd-sourcing. This directly measures the perception of a specific urban environment and could replace the survey part of our method. However, note that the Place Pulse study only measures a few factors (e.g., safety, beauty), limiting the options for streetscape analysis, while the style definition in our study is flexible and not necessarily restricted to health and wellbeing outcomes. Finally, the new method could eventually be used for virtual reality applications, where computer-generated design allows for an immersive and cost-effective evaluation of citywide design interventions, without the need for a potentially slower, more expensive, bottom-up design approach. Note that consecutive images along a single street were rendered in a consistent manner (results not presented).

Overall, this study has utility for researchers, urban planners and urban designers seeking to understand the perception and impact of streetscape design. It may, in the future, directly assist designers and policy-makers to implement our understanding of these relationships using targeted design interventions to improve the health and wellbeing of city residents. Finally, our study provides a tool that could be used in local urban design initiatives to inspire citizen-stakeholders who take an active role in designing their neighbourhood.

Table 2

Average MSE, PSNR (dB) and SSIM statistics between original and translated images. A higher MSE, lower PSNR and lower SSIM indicate less similarity.

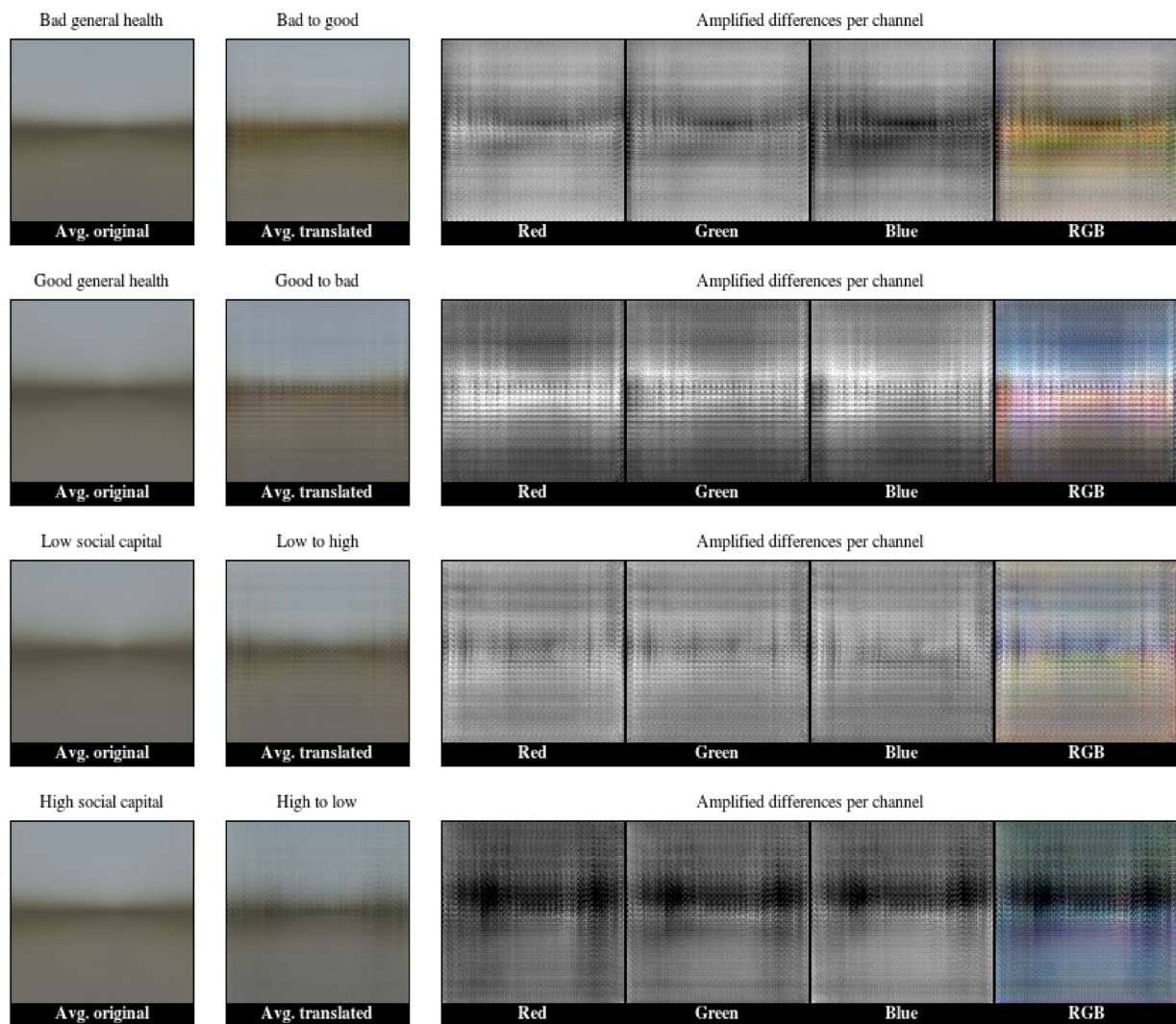|	Low to high			High to low			Average		
	MSE	PSNR	SSIM	MSE	PSNR	SSIM	MSE	PSNR	SSIM
Density	942	19.0	0.57	1151	18.0	0.55	1047	18.5	0.56
Park/city	2191	15.1	0.59	1149	17.9	0.65	1670	16.5	0.62
General health	811	19.6	0.63	1055	18.5	0.61	933	19.0	0.62
Social capital	893	19.0	0.58	876	19.2	0.60	885	19.1	0.59
Life satisfaction	695	20.1	0.65	673	20.4	0.64	684	20.3	0.64

Fig. 8. Average changes the generator neural network makes for different style conversions. For the red, green and blue channel differences, white indicates an increase in that channel, grey (128, 128, 128) no change and black a reduction.

Table 3
Changes in colour channels for the average image per domain.

	Original			Translated			Difference (%)		
	Red	Green	Blue	Red	Green	Blue	Red	Green	Blue
Density – low to high	125	127	121	123	124	118	−1.7	−2.5	−2.1
Density – high to low	123	123	118	128	127	118	4.6	3.4	−0.6
Park to city	114	116	108	94	93	91	−17.8	−19.5	−16.3
City to park	107	108	104	109	111	106	1.8	2.9	1.7
General health – low to high	124	125	120	128	128	119	3.5	2.2	−0.4
General health – high to low	123	124	119	123	124	121	−0.1	−0.3	1.5
Social capital – low to high	121	122	117	124	125	120	3.2	2.6	2.3
Social capital – high to low	124	125	120	114	118	116	−7.3	−5.1	−3.4
Life satisfaction – low to high	124	125	120	128	130	122	3.3	3.8	1.8
Life satisfaction – high to low	122	123	118	117	122	115	−4.0	−0.7	−3.1

Acknowledgements

References

Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., et al. (2010). Google Street View: Capturing the world at street level. *Computer, 43*, 32–38. https://doi.org/10.1109/MC.2010.170.

Bing, B. (2015). *Assessing and Enhancing Video Quality.* Wiley-Blackwell 165–182. https://doi.org/10.1002/9781119133346.ch5 (chapter 5).

Bowler, D. E., Buyung-Ali, L., Knight, T. M., & Pullin, A. S. (2010). Urban greening to cool towns and cities: A systematic review of the empirical evidence. *Landscape and Urban Planning, 97*, 147–155. https://doi.org/10.1016/j.landurbplan.2010.05.006.

Branas, C. C., Kondo, M. C., Murphy, S. M., South, E. C., Polsky, D., & MacDonald, J. M. (2016). Urban blight remediation as a cost-beneficial solution to firearm violence. *American Journal of Public Health, 106*, 2158–2164. https://doi.org/10.2105/AJPH.2016.303434.

Carmona, M., Tiesdell, S., Heath, T., & Oc, T. (2010). *Public Places – Urban Spaces: The Dimensions of Urban Design* (2nd ed.). Elsevier/Architectural Press.

Cohen, D. A., Inagami, S., & Finch, B. (2008). The built environment and collective efficacy. *Health & Place, 14*, 198–208. https://doi.org/10.1016/j.healthplace.2007.06.001.

Curto, J. D., Zarza, I. C., De La Torre, F., King, I., & Lyu, M. R. (2017). *High-resolution deep convolutional generative adversarial networks.* arXiv:1711.06491.

Dantzig, G. B., & Saaty, T. L. (1973). *Compact City: A Plan for a Liveable Urban Environment.* W. H. Freeman.

Department of Health and Human Services (2017). *Victorian Population Health Survey 2015: Selected Survey Findings* State of Victoria Technical Report.

Doersch, C., Singh, S., Gupta, A., Sivic, J., & Efros, A. (2012). What makes Paris look like Paris? *ACM Transactions on Graphics, 31* https://hal.inria.fr/hal-01053876.

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. https://doi.org/10.1007/978-3-319-46448-0_12.

Ewing, R., Hajrasouliha, A., Neckerman, K. M., Purciel-Hill, M., & Greene, W. (2016). Streetscape features related to pedestrian activity. *Journal of Planning Education and Research, 36*, 5–15. https://doi.org/10.1177/0739456X15591585.

Fishman, R. (1982). *Urban Utopias in the Twentieth Century: Ebenezer Howard, Frank Lloyd Wright, and Le Corbusier* (1st ed.). MIT Press.

Foster, S., & Giles-Corti, B. (2008). The built environment, neighborhood crime and constrained physical activity: An exploration of inconsistent findings. *Preventive Medicine, 47*, 241–251. https://doi.org/10.1016/j.ypmed.2008.03.017.

Frank, L. D., & Engelke, P. O. (2001). The built environment and human activity patterns: Exploring the impacts of urban form on public health. *Journal of Planning Literature, 16*, 202–218. https://doi.org/10.1177/08854120122093339.

Frumkin, H. (2002). Urban sprawl and public health. *Public Health Reports, 117*, 201–217. http://www.jstor.org/stable/4598743.

Furuseth, O. J. (1997). Neotraditional planning: A new strategy for building neighborhoods? *Land Use Policy, 14*, 201–213. https://doi.org/10.1016/S0264-8377(97)00002-1.

Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2414–2423). https://doi.org/10.1109/CVPR.2016.265.

Giles-Corti, B., Ryan, K., & Foster, S. (2012). *Increasing Density in Australia: Maximising the Health Benefits and Minimising Harm* National Heart Foundation of Australia Technical Report.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.). *Advances in Neural Information Processing Systems 27* (pp. 2672–2680). Curran Associates, Inc. http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf.

Groenewegen, P. P., van den Berg, A. E., de Vries, S., & Verheij, R. A. (2006). Vitamin G: Effects of green space on health, well-being, and social safety. *BMC Public Health, 6*, 149. https://doi.org/10.1186/1471-2458-6-149.

Hensley, M., Mateo-Babiano, D., & Minnery, J. (2014). Healthy places, active transport and path dependence: A review of the literature. *Health Promotion Journal of Australia, 25*, 196–201. https://doi.org/10.1071/HE14042.

Howard, E. (1902). *Garden cities of To-morrow* (2nd ed.). Swan Sonnenschein & Company, Ltd.

Huang, X., Liu, M. Y., Belongie, S., & Kautz, J. (2018). *Multimodal unsupervised image-to-image translation.* arXiv:1804.04732.

ImageMagick Studio LLC (2018). *ImageMagick.* , version 7 https://www.imagemagick.org.

International Organization for Migration (2015). *World Migration Report 2015 – Migrants and Cities: New Partnerships to Manage Mobility* IOM Technical Report.

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 5967–5976). https://doi.org/10.1109/CVPR.2017.632.

Jabareen, Y. R. (2006). Sustainable urban forms: Their typologies, models, and concepts. *Journal of Planning Education and Research, 26*, 38–52. https://doi.org/10.1177/0739456X05285119.

Joss, S., Cowley, R., & Tomozeiu, D. (2013). Towards the 'ubiquitous eco-city': An analysis of the internationalisation of eco-city policy and practice. *Urban Research & Practice, 6*, 54–74. https://doi.org/10.1080/17535069.2012.762216.

Kingma, D. P., & Welling, M. (2013). *Auto-encoding variational Bayes.* arXiv:1312.6114.

Kingsley, J. Y., & Townsend, M. (2006). 'Dig In' to social capital: Community gardens as mechanisms for growing urban social connectedness. *Urban Policy and Research, 24*, 525–537. https://doi.org/10.1080/08111140601035200.

Kondo, M. C., South, E. C., & Branas, C. C. (2015). Nature-based strategies for improving urban health and safety. *Journal of Urban Health, 92*, 800–814. https://doi.org/10.1007/s11524-015-9983-y.

Le Corbusier (1933). *La ville radieuse*.

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 105–114). . https://doi.org/10.1109/CVPR.2017.19.

Leyden, K. M. (2003). Social capital and the built environment: The importance of walkable neighborhoods. *American Journal of Public Health, 93*, 1546–1551. https://doi.org/10.2105/AJPH.93.9.1546.

Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening, 14*, 675–685. https://doi.org/10.1016/j.ufug.2015.06.006.

Liu, L., Zhou, B., Zhao, J., & Ryan, B. D. (2016). C-IMAGE: City cognitive mapping through geo-tagged photos. *GeoJournal, 81*, 817–861. https://doi.org/10.1007/s10708-016-9739-6.

Liu, M. Y., Breuel, T., & Kautz, J. (2017). Unsupervised image-to-image translation networks. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.). *Advances in Neural Information Processing Systems 30* (pp. 700–708). Curran Associates, Inc. http://papers.nips.cc/paper/6672-unsupervised-image-to-image-translation-networks.pdf.

Lozano, E. E. (1990). *Community Design and the Culture of Cities: The Crossroad and the Wall* (1st ed.). Cambridge University Press.

Lynch, K. (1960). *The Image of the City.* MIT Press.

Maas, J., van Dillen, S. M. E., Verheij, R. A., & Groenewegen, P. P. (2009). Social contacts as a possible mechanism behind the relation between green space and health. *Health & Place, 15*, 586–595. https://doi.org/10.1016/j.healthplace.2008.09.006.

Maas, J., Verheij, R. A., Groenewegen, P. P., de Vries, S., & Spreeuwenberg, P. (2006). Green space, urbanity, and health: How strong is the relation? *Journal of Epidemiology & Community Health, 60*, 587–592. https://doi.org/10.1136/jech.2005.043125.

Mair, C., Diez Roux, A. V., & Galea, S. (2008). Are neighbourhood characteristics associated with depressive symptoms? A review of evidence. *Journal of Epidemiology & Community Health, 62*, 940–946. https://doi.org/10.1136/jech.2007.066605.

Mehaffy, M. W., Porta, S., & Romice, O. (2015). The "neighborhood unit" on trial: A case study in the impacts of urban morphology. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability, 8*, 199–217. https://doi.org/10.1080/17549175.2014.908786.

Naik, N., Philipoom, J., Raskar, R., & Hidalgo, C. (2014). Streetscore – Predicting the perceived safety of one million streetscapes. *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 793–799). https://doi.org/10.1109/CVPRW.2014.121.

Nguyen, D. (2010). Evidence of the impacts of urban sprawl on social capital. *Environment and Planning B: Planning and Design, 37*, 610–627. https://doi.org/10.1068/b35120.

Nice, K. A., Wijnands, J. S. (2019). Melbourne Google Street View imagery dataset. Data in Brief.

Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T., & Efros, A. A. (2016). Context encoders: Feature learning by inpainting. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2536–2544). https://doi.org/10.1109/CVPR.2016.278.

Pereira, G., Foster, S., Martin, K., Christian, H., Boruff, B. J., Knuiman, M., et al. (2012). The association between neighborhood greenness and cardiovascular disease: An observational study. *BMC Public Health, 12*, 1–9. https://doi.org/10.1186/1471-2458-12-466.

Perry, C. A. (1929). *The neighborhood unit, a scheme of arrangement for the family-life community. Regional Plan of New York and its Environs, vol. 7.*

PSMA Australia Ltd (2017). *PSMA street network (August 2017).* accessed from AURIN Portal on 28 February 2018.

QGIS Development Team (2018). *QGIS Geographic Information System. Open source geospatial foundation project.* , version 2.18.17 https://www.qgis.org.

Rutheiser, C. (1997). Beyond the radiant garden city beautiful: Notes on the new urbanism. *City & Society, 9*, 117–133. https://doi.org/10.1525/ciso.1997.9.1.117.

Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The collaborative image of the city: Mapping the inequality of urban perception. *PLOS ONE, 8*, 1–12. https://doi.org/10.1371/journal.pone.0068400.

Sallis, J. F., Cain, K. L., Conway, T. L., Gavand, K. A., Millstein, R. A., Geremia, C. M., et al. (2015). Is your neighborhood designed to support physical activity? A brief streetscape audit tool. *Preventing Chronic Disease, 12*, 1–11. https://doi.org/10.5888/pcd12.150098.

Seiferling, I., Naik, N., Ratti, C., & Proulx, R. (2017). Green streets – Quantifying and mapping urban trees with street-level imagery and computer vision. *Landscape and Urban Planning, 165*, 93–101. https://doi.org/10.1016/j.landurbplan.2017.05.010.

Sharifi, A. (2016). From Garden City to Eco-urbanism: The quest for sustainable neighborhood development. *Sustainable Cities and Society, 20*, 1–16. https://doi.org/10.1016/j.scs.2015.09.002.

Smith, C., & Levermore, G. (2008). Designing urban spaces and buildings to improve sustainability and quality of life in a warmer world. *Energy Policy, 36*, 4558–4562. https://doi.org/10.1016/j.enpol.2008.09.011.

Sønderby, C. K., Caballero, J., Theis, L., Shi, W., & Huszár, F. (2017). Amortised MAP inference for image super-resolution. *5th International Conference on Learning Representations*, 1–17.

South, E. C., Hohl, B. C., Kondo, M. C., MacDonald, J. M., & Branas, C. C. (2018). Effect of greening vacant land on mental health of community-dwelling adults: A cluster randomized trial. *JAMA Network Open, 1*, e180298. https://doi.org/10.1001/jamanetworkopen.2018.0298.

Spokane, A. R., Lombard, J. L., Martinez, F., Mason, C. A., Gorman-Smith, D., Plater-Zyberk, E., et al. (2007). Identifying streetscape features significant to well-being. *Architectural Science Review, 50*, 234–245. https://doi.org/10.3763/asre.2007.5029.

Stevenson, M., Thompson, J., de Sá, T. H., Ewing, R., Mohan, D., McClure, R., et al. (2016). Land use, transport, and population health: Estimating the health benefits of compact cities. *The Lancet, 388*, 2925–2935. https://doi.org/10.1016/S0140-6736(16)30067-8.

Stradling, S. G., Anable, J., & Carreno, M. (2007). Performance, importance and user disgruntlement: A six-step method for measuring satisfaction with travel modes. *Transportation Research Part A: Policy and Practice, 41*, 98–106. https://doi.org/10.1016/j.tra.2006.05.013.

Sugiyama, T., Leslie, E., Giles-Corti, B., & Owen, N. (2008). Associations of neighbourhood greenness with physical and mental health: Do walking, social coherence and local social interaction explain the relationships? *Journal of Epidemiology & Community Health, 62*, 1–6. https://doi.org/10.1136/jech.2007.064287.

Sullivan, W. C., Kuo, F. E., & Depooter, S. F. (2004). The fruit of urban nature: Vital neighborhood spaces. *Environment and Behavior, 36*, 678–700. https://doi.org/10.1177/0193841X04264945.

Tsolakis, N., & Anthopoulos, L. (2015). Eco-cities: An integrated system dynamics framework and a concise research taxonomy. *Sustainable Cities and Society, 17*, 1–14. https://doi.org/10.1016/j.scs.2015.03.002.

Turrell, G., Haynes, M., Wilson, L. A., & Giles-Corti, B. (2013). Can the built environment reduce health inequalities? A study of neighbourhood socioeconomic disadvantage and walking for transport. *Health & Place, 19*, 89–98. https://doi.org/10.1016/j.

healthplace.2012.10.008.

Tzoulas, K., Korpela, K., Venn, S., Yli-Pelkonen, V., Kaźmierczak, A., Niemela, J., et al. (2007). Promoting ecosystem and human health in urban areas using green infrastructure: A literature review. *Landscape and Urban Planning, 81*, 167–178. https://doi.org/10.1016/j.landurbplan.2007.02.001.

de Vries, S., van Dillen, S. M. E., Groenewegen, P. P., & Spreeuwenberg, P. (2013). Streetscape greenery and health: Stress, social cohesion and physical activity as mediators. *Social Science & Medicine, 94*, 26–33. https://doi.org/10.1016/j.socscimed.2013.06.030.

de Vries, S., Verheij, R. A., Groenewegen, P. P., & Spreeuwenberg, P. (2003). Natural environments–healthy environments? An exploratory analysis of the relationship between greenspace and health. *Environment and Planning A: Economy and Space, 35*, 1717–1731. https://doi.org/10.1068/a35111.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing, 13*, 600–612. https://doi.org/10.1109/TIP.2003.819861.

Ward Thompson, C., Roe, J., Aspinall, P., Mitchell, R., Clow, A., & Miller, D. (2012). More green space is linked to less stress in deprived communities: Evidence from salivary cortisol patterns. *Landscape and Urban Planning, 105*, 221–229. https://doi.org/10.1016/j.landurbplan.2011.12.015.

Wright, F. L. (1932). *The Disappearing City* (1st ed.). William Farquahar Payson.

Yeh, R. A., Chen, C., Lim, T. Y., Schwing, A. G., Hasegawa-Johnson, M., & Do, M. N. (2017). *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 6882–6890). https://doi.org/10.1109/CVPR.2017.728.

Yin, L., & Wang, Z. (2016). Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. *Applied Geography, 76*, 147–153. https://doi.org/10.1016/j.apgeog.2016.09.024.

Zhou, B., Liu, L., Oliva, A., & Torralba, A. (2014). Recognizing city identity via attribute analysis of geo-tagged images. *Computer Vision – ECCV 2014*, 519–534 Springer.

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017a). Unpaired image-to-image translation using cycle-consistent adversarial networks. *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 2242–2251). https://doi.org/10.1109/ICCV.2017.244.

Zhu, J. Y., Zhang, R., Pathak, D., Darrell, T., Efros, A. A., Wang, O., et al. (2017b). Toward multimodal image-to-image translation. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.). *Advances in Neural Information Processing Systems 30* (pp. 465–476). Curran Associates, Inc. http://papersnips.cc/paper/6650-toward-multimodal-image-to-image-translation.pdf.