PLACEMAKINGAI:
PARTICIPATORY URBAN DESIGN WITH GENERATIVE
ADVERSARIAL NETWORKS

DONGYUN KIM¹, GEORGE GUIDA² and JOSE LUIS GARCÍA
DEL CASTILLO Y LÓPEZ ³
¹,²,³Harvard University Graduate School of Design.
¹dongyun_kim@gsd.harvard.edu, 0000-0003-1193-3845
²gguida@gsd.harvard.edu, 0000-0002-2477-0064
³jgarciadelcasti@gsd.harvard.edu, 0000-0001-6117-1602

Abstract. Machine Learning (ML) is increasingly present within the
architectural discipline, expanding the current possibilities of
procedural computer-aided design processes. Practical 2D design
applications used within concept design stages are however limited by
the thresholds of entry, output image fidelity, and designer agency. This
research proposes to challenge these limitations within the context of
urban planning and make the design processes accessible and
collaborative for all urban stakeholders. We present PlacemakingAI, a
design tool made to envision sustainable urban spaces. By converging
supervised and unsupervised Generative Adversarial Networks (GANs)
with a real-time user interface, the decision-making process of planning
future urban spaces can be facilitated. Several metrics of walkability
can be extracted from curated Google Street View (GSV) datasets when
overlayed on existing street images. The contribution of this framework
is a shift away from traditional design and visualization processes,
towards a model where multiple design solutions can be rapidly
visualized as synthetic images and iteratively manipulated by users. In
this paper, we discuss the convergence of both a generative image
methodology and this real-time urban prototyping and visualization
tool, ultimately fostering engagement within the urban design process
for citizens, designers, and stakeholders alike.

Keywords. Machine Learning; Generative Adversarial Networks;
User Interface; Real-time; Walkability; SDG 11.

1. Introduction
The process of urban planning is an iterative and collaborative process, often limited
to designers and city stakeholders. There is an inherent challenge in communicating
and inclusively producing design solutions. If capitalizing on GANs (Goodfellow et
al., 2014) can aggregate hundreds of images and characteristics of spaces, an intuitive
design tool can help bridge this missing communication. With the intention to imagine
and conceive improved future cities, this project speculates on ways in which

POST-CARBON, Proceedings of the 27th International Conference of the Association for Computer-
Aided Architectural Design Research in Asia (CAADRIA) 2022, Volume 2, 485-494. © 2022 and
published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA),
Hong Kong.
computational statistical models can facilitate this dialogue from citizens and architects to urban planners and policymakers.

In this paper, we present PlacemakingAI, a tool that allows users to envision improved streetscapes and alternative versions of existing cities. We introduce methods of collecting datasets based on metrics of walkability, designed to train on two known GAN models – CycleGAN (Zhu et al., 2017) and Pix2Pix (Isola et al., 2016), to generate synthetic images in real-time. By encoding the concepts of genius loci of space (Norber-Schulz, 1980) – in particular, the character, life, and the everyday, and imageability of space (Lynch, 1960) – in this case, the node or street as a connector and way-finder, the resultant images can provide insights of novel and improved urban spaces. By explaining this framework, we will show a feasibility study and describe a real-time participatory user interface (UI) that could be used by all urban stakeholders and conclude with limitations and future outlooks.

Figure 1. Generated street scene with embedded metrics of urban walkability.

2. Background: Machine Learning and the Perception of the City

The concept of imageability by Kevin Lynch finds relevance to this day in urban design and city planning practices. First introduced in his ‘Image of the City’, it presents a novel way of reading and observing the city based on the classification of paths, edges, districts, nodes, and landmarks. The aggregation of these elements forms the memory or mental images of place, which informs this paper’s methodology in the participatory creation of images of improved green and public spaces. Encompassing this within a real-time UI gives agency to citizens and designers to improve collaborative urban planning.

Of Lynch’s 5 major elements in cities, the path can be considered a space of most active human experience, including streets, sidewalks, trails, and canals (Lynch, 1960). Functioning as the binding component that connects all other elements, this will be the main element considered when visualizing future sustainable cities. Norber-Schultz’s reading of the city instead provides a granular understanding of place. This environment is composed firstly of concrete things, having material substance, shape,
PLACEMAKINGAI:
PARTICIPATORY URBAN DESIGN WITH GENERATIVE ADVERSARIAL NETWORKS

texture, color, and most importantly character. Defined by the phenomenology of everyday life, the character is then defined by how things are, providing insight into the genius loci or “spirit of a place” (Norber-Schulz, 1980) as first identified by Kyle Steinfield in Gan Loci (Steinfield, 2019).

The phenomenological reading of the city and its potential in visualizing the potential future genius loci of place through machine learning encounters additional practical applications this paper seeks to address. With the recent shortcomings of many cities during the COVID-19 pandemic, it has become increasingly meaningful to maintain a high quality of the environment in urban areas. Home to around 40% of the EU population, the average satisfaction of green and public spaces is at 77%, however, in cities such as Athens the overall satisfaction of public and green spaces remains critically low (Dijkstra et al., 2020). Interactive tools such as Treepedia (Li et al., 2015) provide awareness on the urban vegetation present in cities through a Green View Index, which indicates how several large cities have contrasting levels of green spaces indices ranging from Paris (8.8%) to Singapore (29.3%).

From the perspective of the seventeen UN Sustainable Development Goals, this project seeks to “enhance inclusive and sustainable urbanization and capacity for participatory, integrated and sustainable human settlement planning and management in all countries” in the creation of “safe, inclusive, and accessible green and public spaces” for all citizens; GOAL 11: Sustainable Cities and Communities (United Nations, 2015). With this, the work seeks to expand on current research on the quantitative assessment of qualitative urban data (Zhang, et al., 2018) into an accessible generative space planning tool. This is made possible by identifying a set of spatial characteristics based on metrics of walkability - a term that can be viewed from both qualitative and quantitative perspectives. Lynch’s concept of imageability can be used to identify physical features and urban design qualities such as enclosure, human scale, transparency, linkage, complexity, and coherence (Ewing and Handy, 2009). The genius loci can be used to add a subjective reading of an urban environment and its atmosphere. This can include measures of individuals’ senses of safety, comfort, and
levels of interest, further establishing a relationship between the physical features of a street and walking user behavior (Ewing and Handy, 2009). Jan Gahl describes this as “life in the space, the climate, and the architectural quality support and complement each other to create an unforgettable total impression” (Gehl, 2011). The subjectivity associated with walkability presents wide opportunities to reimagine the city while confirming an implicit bias during the design ideation and evaluation process.

Previous architectural and urban design applications of ML have ranged from stylistically designing 2D plans (Del Campo et al., 2019), and building facades (Özel & Ennemoser 2020), to form-based optimization with SpacemakerAI (Spacemaker, n.d.). Each of these embraces human-machine collaboration to different degrees while acknowledging the limitations of human agency within the design process. Due to the inherently iterative and collaborative process of urban planning, extending the feedback loop of the user beyond that of the collection of the dataset into a comprehensive UI becomes critical. The combination of tangible UI’s and web-based platforms such as CityScope from the MIT Media Lab (Noyman et al., 2018) has improved this lacking dialogue. The project ambition is therefore to speculate the use of this real-time prototyping GAN tool to be used as a platform for citizens, designers, and city stakeholders to have meaningful conversations on improving the environment in urban streets.

3. Methodology

PlacemakingAI is a design tool intended to provide ease in creating iterative design solutions and supporting user feedback. This UI prototype facilitates the use of machine learning models to stylize images or 3D models of city streets, based on neural networks trained on walkable street datasets. In this section, we describe the main steps in the development of this prototype including Data preparation, Model Training, and the real-time UI.

Figure 3. PlacemakingAI concept workflow and feedback loop.
3.1. DATA PREPARATION

The first step in the process of training each GAN model is the selection and curation of the image datasets. While the term walkability adopts a broad spectrum of metrics, we first established the general criteria and then divided these into two proof of concept datasets for users to choose from. The first step was to select cities and streetscape images that satisfy a variety of urban design qualities relating to walkability. These include qualitative qualities such as enclosure, human scale, transparency, linkage, complexity, coherence (Ewing & Handy, 2009), and that ideally contain a lively atmosphere. The second step was to manually separate these into Dataset A: Green Spaces, which contains active green street scenes with trees, plants, and low-level greenery, and Dataset B: Pedestrian Commercial, which contains lively pedestrian-only streets, with commercial frontages, kiosks, restaurants, or cafes. Both datasets comprised of 620 images, which were captured from Google Street View within Google Earth at resolutions above 1024x1024 pixels. Images were selected from model streets and cities across North, West, and Southern European cities familiar to the authors (Table 1). The data was then augmented, a common practice in machine learning to improve output resolution, to 3720 images by extracting the left, center, and right quadrant of each image and additionally mirroring these.

<table>
<thead>
<tr>
<th>Street</th>
<th>City</th>
<th>Dataset A</th>
<th>Dataset B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrer de Gràcia</td>
<td>Barcelona</td>
<td></td>
<td></td>
</tr>
<tr>
<td>La Rambla</td>
<td>Barcelona</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paseig del Born</td>
<td>Barcelona</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paseo del Prado</td>
<td>Madrid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calle de Alcalá</td>
<td>Madrid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rambla del Raval</td>
<td>Barcelona</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passeig de Gràcia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passeig de Gràcia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carrer de Gràcia</td>
<td>Barcelona</td>
<td></td>
<td></td>
</tr>
<tr>
<td>La Rambla</td>
<td>Barcelona</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paseig del Born</td>
<td>Barcelona</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paseo del Prado</td>
<td>Madrid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calle de Alcalá</td>
<td>Madrid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rambla del Raval</td>
<td>Barcelona</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passeig de Gràcia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passeig de Gràcia</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.2. MODEL TRAINING

In this research, we chose supervised and unsupervised machine learning GAN models to explore their respective applications within the architecture discipline.

The first was Pix2Pix, a commonly used supervised machine learning model composed of a conditional GAN. We chose to use this model specifically to reconstruct
synthetic images of street scenes based on labeled or segmented images, for the most effective results. By labeling street images, we can extract several key features that define each street and allow us to implant these characteristics onto the targeted street, ranging from roads, building frontages, people, benches, trees, etc. This process requires two inputs to train the Pix2Pix model: the original street image (Image A) and the equivalent segmented image (Image B). Therefore, Dataset A was trained with segmented Dataset A, and Dataset B was trained with segmented Dataset B (Figure 4). The image segmentation was implemented in PyTorch on MIT’s ADE20K (Zhou et al., 2018), one of the largest open-source datasets for semantic segmentation and scene parsing. 1024x1024 pixel segmented images were created of both datasets. The second model, CycleGAN is in contrast based on an unsupervised GAN architecture, allowing the use of an unpaired dataset. Therefore, each dataset did not require the additional segmented dataset. Both models were equally run for 200 epochs of training cycles over the course of 2 days, on Google Colab Pro environment (P100 GPU was used).

For the training of both models, a new Dataset C was introduced, based on 3D viewport images of streets from a selected test city. Both Datasets A and B were trained separately (in each model) with Dataset C (Figure 4) to provide users with multiple trained models to select from. A 3D digital model of a test city with streetscapes, in this case, Cambridge, USA, was used to provide greater user control over the generated output within the user interface. This same methodology could be used with an image overlay, for example of a derelict, misused, or car-centric street. However, this would limit the user control and output image fidelity as explained in the following section.

4. Results & Discussions

4.1. TRAINING RESULTS

Both models present diverse outcomes, reflecting on their supervised and unsupervised architecture. In the Pix2Pix model, the results trained from the green Dataset A revealed a variety of trees and greenery, while the pedestrian model indicates a lively urban character, containing amenities and people. This spatial identity could be used to
determine and visually evaluate the quality of space.

The outcomes of both training models produced varied results, representing a diverse synthesis of the different European streets. Both formal qualities and atmosphere or character (genius loci) emerge to different degrees and are intended to be taken as a starting point within the current exponential developments in machine learning. Several iterations demonstrated how increases in the dataset size and quality generated greener, livelier, and more varied streetscapes.

Figure 5. Sample output images.

Table 2. Training result comparison.
The results demonstrated by the Pix2Pix model show notably better results than CycleGAN, capturing specific spatial features to a greater extent. Figure 5 shows that certain synthetic output images begin resembling the abstracted white facades of Oxford Street, street activity of La Rambla, or bright green trees in the streets of Amsterdam.

This visual difference in part emerges from the difference between supervised and unsupervised learning. For this reason, Pix2Pix can be used to generate more detailed results per segmented object, while CycleGAN can be utilized to get unexpected results, at times mapping building surfaces with green textures.

4.2. REAL-TIME USER INTERFACE

The UI is intended for both citizens and stakeholders alike to visualize and interact in real-time with selected street views. Prototyped on HumanUI (Human UI, 2016), a plugin to Rhinoceros Grasshopper, several user inputs and sliders were used to provide high levels of design control (Figure 6). The inputs are divided into two parts. The first includes the selection of one of the two trained models, Datasets A and B, and the selection of a camera location based on the chosen city. The second is related to the urban inputs within the 3D scene. These include the manipulation of sample urban elements including pedestrian, tree, and bench count or pavement width, which visually improve the streetscape quality of the generated images.

This novel method of using GANs reflects a similar way designers originate ideas. In the same way, one sketches or uses architectural software, a feedback loop and level of the agency are provided, revealing future potentials of AI. On the left display, users
can interact seamlessly with their chosen street in 3D and design intent, and on the right screen, the chosen machine learning model presents the generated output in real-time. Users can zoom in, pan, and orbit around the scene to the preferred camera location on a chosen street.

In the back end, the current Rhino viewport was refreshed every 20 frames per second and this image was directly fed through custom python components into the Pix2Pix model to generate the resulting image. Each time a new model, dataset, camera or scene manipulation was made, both images would update. This however revealed the limitations in computing power to process high-resolution images, for example of 1024x1024 pixels. While this methodology was successful, only 256x256 pixel images provided sufficient real-time updates to respond to the user inputs on Human UI.

5. Conclusion

PlacemakingAI explores the possibilities of GANs in the generation and visualization of walkable urban streets. Building on the concept of imageability and genius loci, physical features, urban design qualities, and atmospheric characteristics of space can become embedded within the trained machine learning models. The variety of images produced and visualized provides an insight into early developments of real-time ML applications within the architecture engineering and construction industry and how generating and communicating design within the urban design could be overcome. This tool provides a framework in response to the UN Sustainability Goal 11: Sustainable Cities and Communities in enhancing this participatory process of urbanization into improved public spaces. Future work seeks to expand on the notion of feedback and accessibility, providing users with further intuitive design and dataset creation tools for increased agency and user testing through web-based applications or tangible user interfaces. PlacemakingAI seeks to advance current uses of machine learning into participatory processes and human-centric design solutions.

References


