Enhancing Urban Design Through Geodata and Machine Learning

Alex Mackay and Diego Pajarito Grajales

Urban Big Data Centre, University of Glasgow
January 20, 2023

Summary

Machine learning (ML) methods have seen surprisingly little application in the innovation-driven field of urban design. Research by Boim, Dortheimer, and Sprecher (2022) presented a novel use of ML to generate alternative urban plans considerate of existing local practices, using data extracted from a GIS package to train a Conditional Generative Adversarial Network (CGAN) model. This paper extends that work with a novel dataset built from open geospatial data from Glasgow with results validated using additional quantitative, and qualitative validation methods. The results show the CGAN model is capable of producing geographic and contextually sympathetic urban design proposals with output quality confirmed by result validation.

KEYWORDS: Urban Design, GIS, Machine Learning, Conditional Generative Adversarial Networks, Pix2PixHD

1. Introduction

Machine learning (ML) as a branch of artificial intelligence (AI) is the algorithmic powerhouse behind many aspects of modern life. However, Urban design, known for its inextricable connection with geographic spaces, technology and innovation, remains estranged from applied artificial intelligence both academically and in practice (Khean, Fabri, Heusler, 2018; Belam, Santos, Leitao, 2019). This is due in part to the novelty of machine learning methods, and partly the underlying skills gap surrounding them (Raman, Kollar, Penman, 2022). Most progressive discourse on ML methods has been framed within the fields of computer science, engineering, and geography, with urban researchers only just beginning to engage with the methods.

The main question preoccupying this paper is *'How might machine learning methods in conjunction with geoinformation sciences be used to support urban design?*'. Image-to-image translation using CGAN models for image infilling was identified as a very promising approach in a review of the literature. A CGAN model was also chosen for its simplicity, requiring little coding knowledge, and for its use of image data, a common data type used in design conceptualisation.

2. Data and methods

To prepare a model capable of producing geographic and contextually sympathetic urban design proposals, the authors prepared the GlasGAN site model. It uses a paired dataset of urban map tile images, as done by Boim, Dortheimer and Sprecher (2022) but with data taken from Glasgow's open geospatial data repositories (i.e., DigiMaps OS Maps). Three categories of data are used in this study (see figure 1): Ground truth images 'B', from the geospatial visualization of open data layers built using GIS, are full unmasked images. Input label images 'A', are masked ground truth images. Finally, synthesized images that contains the infilled representations generated by the CGAN.

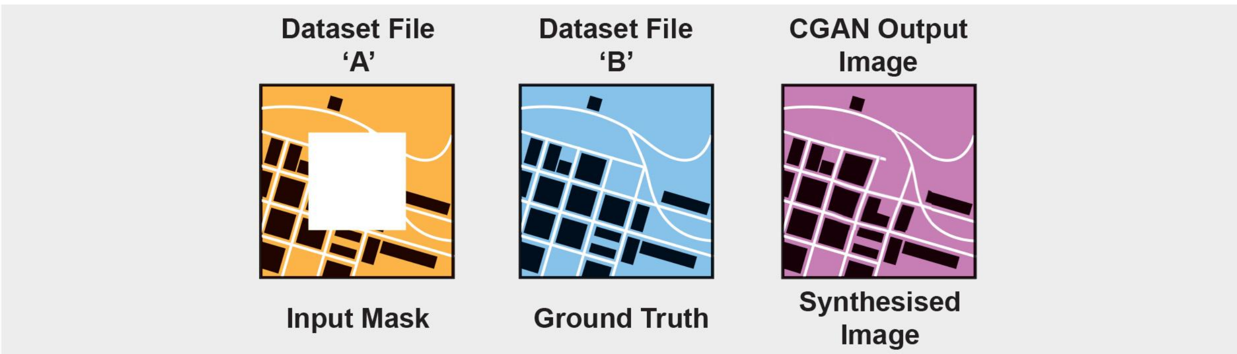


Figure 1 - Dataset structure, showing input mask and ground truth.

2.1 Data

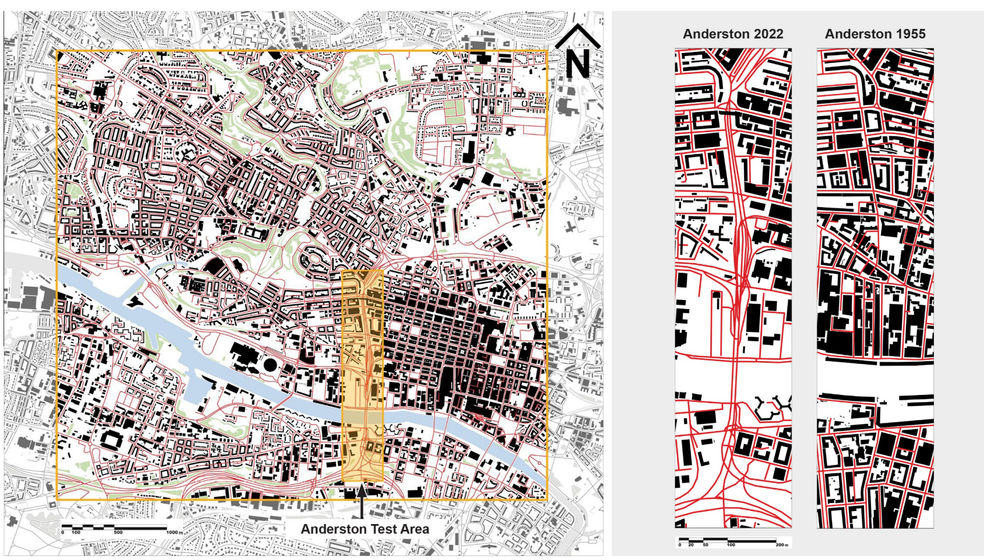


Figure 2 - Area covered by GlasGAN: Site datasets with Anderston test area highlighted.

A base map of Glasgow showing building footprints, roads, tidal water, and green spaces was constructed using DigiMaps OS Maps geospatial data. The four layers were set to resemble Boim, Dortheimer, and Sprecher's simple block colors and lines using ArcGIS Pro. Once assembled, map tiles were automatically exported from the basic map using a Python script developed by Dortheimer (2022). The script produced two sets of 900 images, one of which masked 25% of the canvas with a white square, as seen in most CGAN datasets.

The test dataset, paired images not used during model training, was collected from the Anderston area of the city. The Anderston neighborhood is a key example of urban transformation having been entirely redeveloped in the mid 1950s to make way for the city's M8 motorway. As Anderston sits within the center of the training data area, the typology and design of the test data should be comparable with the modern training data.

The exported dataset was expanded by copying and mirroring the initial 1800 images adding them to the original A and B data sets. Finally, all Pix2PixHD input data needs to be scaled to dimensions with a base of 32 pixels which meant that all 3600 images were resized to 512 x 512 pixels

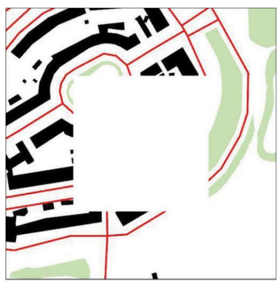
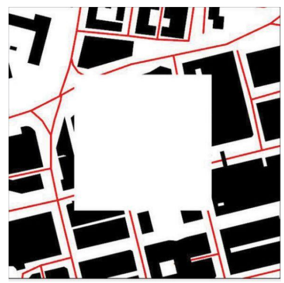
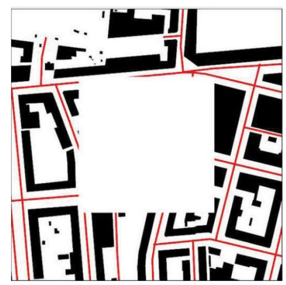
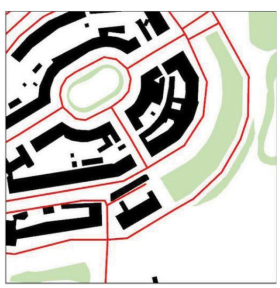
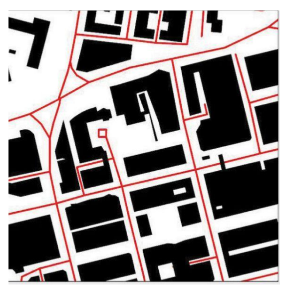
	Dataset 1	**Dataset 2**	**Dataset 2 - Test**
Dataset Contents	Houses, roads, rivers, and green spaces.	Houses and roads.	Houses and roads.
Area Extracted	2022 Glasgow	2022 Glasgow	1955 Anderston
Input Mask 'A'			
Ground Truth 'B'			

Table 1 - The three datasets created for the GlasGAN: Site model.

2.2 Methods

GAN architecture consists of two deep networks: a generator that creates an output and a discriminator that compares it to the expected result (Boim, Dortheimer, Sprecher, 2022). The process of creating images and applying the discriminator is repeated until the discriminator is 'convinced' by the quality result, which strengthens the neural pathways in the model and allows the generator to provide consistent outputs. CGAN models are a modified GAN that receive an input which guides output generation. In this study, the conditional input was the masked test data for the Site model. Pix2PixHD CGAN was chosen for this study for its previous application by Boim, Dortheimer and Sprecher, as well as its ease of use. Pix2PixHD uses the open source Python coding language and supports training on custom datasets.

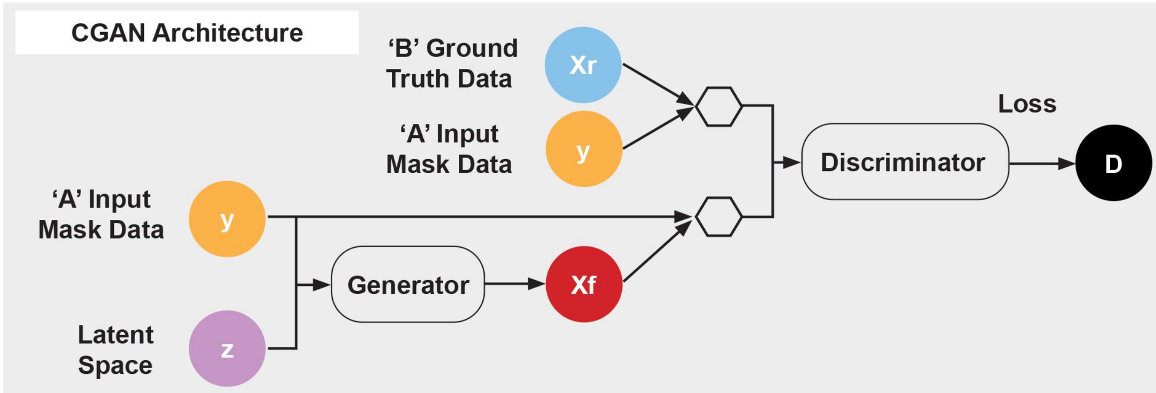


Figure 3 - The architecture of a CGAN model.

2.2.1 CGAN Training and Testing

Pix2PixHD was hosted in a Google Colab notebook with datasets (123Mb approx) uploaded to Google Drive. The source code for Pix2PixHD was cloned from a git repository[1] and the notebook's code ran in a dedicated pool instances to ensure connection during the execution . The Python tool Dominate was then installed to handle CGAN's training and testing HTML files. In total, two separate models were trained, one for each dataset, with one passed on to testing.

2.2.2 Quantitative Output Validation

Quantitative validation was first used to statistically examine the model training results. Since most of the AI/ML projects use losses functions to evaluate models performance, the authors decided to plot the model training losses up to epoch 400. Such plots indicate changes in model performance through additional epochs, as well as whether a model is overfitted.

Structural Similarity Index (SSIM) test was also used to compare the ground truth and synthesized outputs of the 5 control test images. SSIM is a perceptual metric that quantifies image quality degradation and is used to validate image-synthesis outputs, where closer similarity indicates better output quality and accuracy, thus better model training (Salehi and Chalechale, 2020; Qu et al, 2019). The test gives values between 0 and 1, with 1 indicating a perfect image copy. The images were cropped for this test, only showing the central generated portion. Including the full image would have given an unrealistically high SSIM value because it was not masked in training. The test used the Scikit-Image package in Python in the Colab environment.

[1] "Project Name" 's source code is available as an "BSD License" project at: https://github.com/NVIDIA/pix2pixHD

2.2.3 Expert-Driven Output Validation

A survey of output images was used to determine whether synthesized test outputs appeared at a similar quality to control test outputs. Each survey showed 10 synthesized test images, consisting of 5 control test images from training data, and 5 novel test inputs. A number generator was used to avoid selection bias while selecting images. Each survey presented experts one image from the training or test datasets at a time and asked questions (Table 2) about image content.

The surveys were distributed to 11 urban design professionals working closely in the field. Such selection speeded up the validation and ensured a minimal number of surveys to identify initial patterns. Likert scale survey questions were hosted on the Qualtrics experience management website for anonymous data collecting and safe data storage; no personal or identifying information was requested, and all replies were anonymised.

Table 2 - Survey Questions.

Question	In the generated area:
1.	The image is clearly identifiable as an urban map.
2.	The buildings are arranged logically in this space.
3.	The buildings look correctly shaped.
4.	The roads are arranged logically in this space.

Results data for each question were then validated for normal distribution using the Shaprio test, with a T-Test then performed to check for statistical significance between answers for control and test images.

3. Results and Discussion

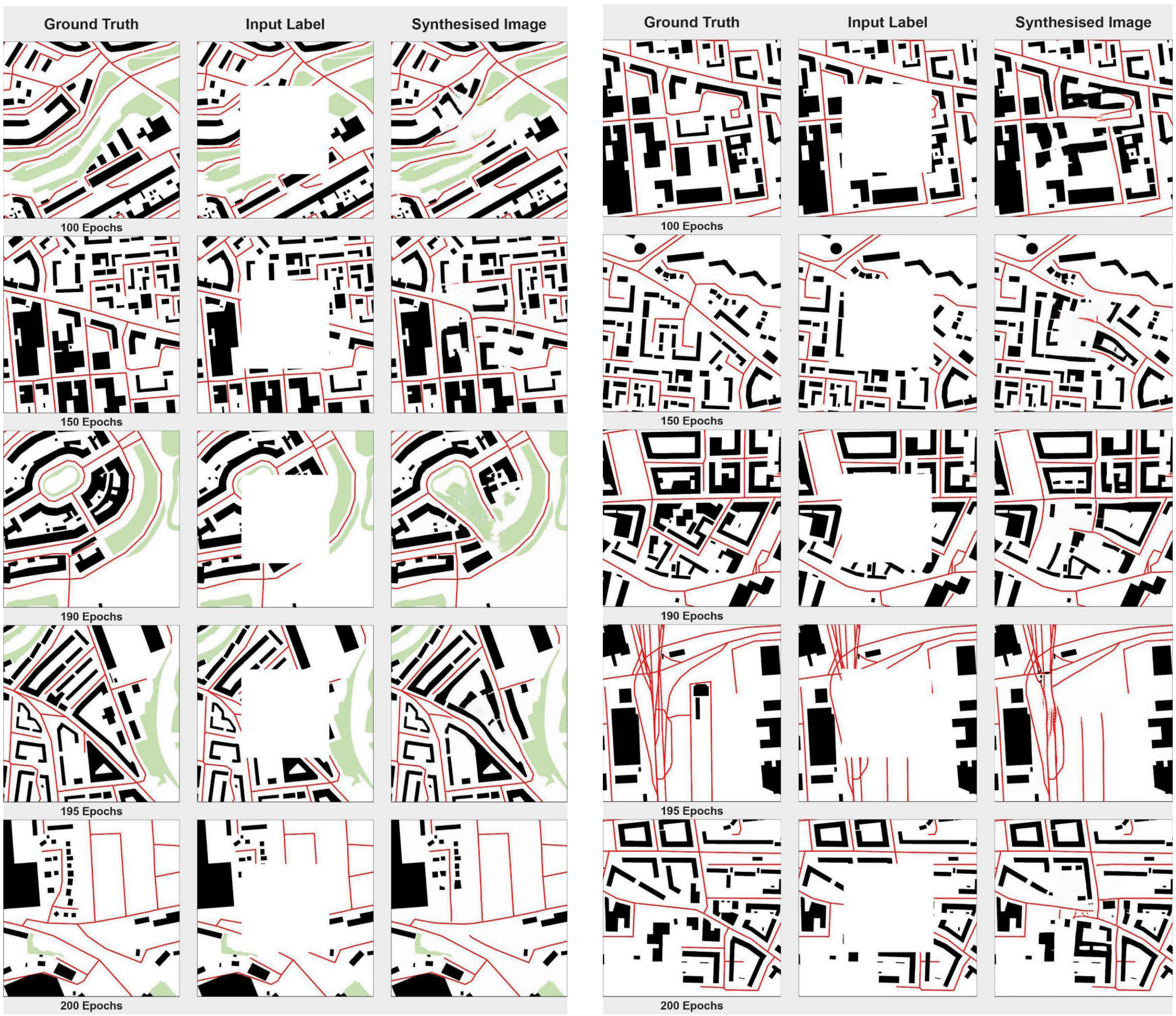


Figure 4 - Sample of training results for GlasGAN model across various epochs.

Two GlasGAN models were trained for 200 epochs respectively. Training results visually showed considerable organic artifacts on Model 1 results. Without natural features, Model 2 outputs showed more consistent improvement with each epoch, better structuring buildings and roads and having fewer image artifacts. Model 2 was chosen to be trained up to epoch 400 and passed to testing.

3.1 Testing

Figure 5 - Sample of control and Test results from GlasGAN: Site Model 2.

After 400 epochs, Model 2 was input with 20 test images. 5 images randomly selected from the training dataset as control images alongside 15 images from the test datasets. The input mask was processed by the model, with the CGAN infilling the missing urban area based on training results. This produced 20 novel output images showing possible urban layouts within the masked area.

3.2 Quantitative Validation

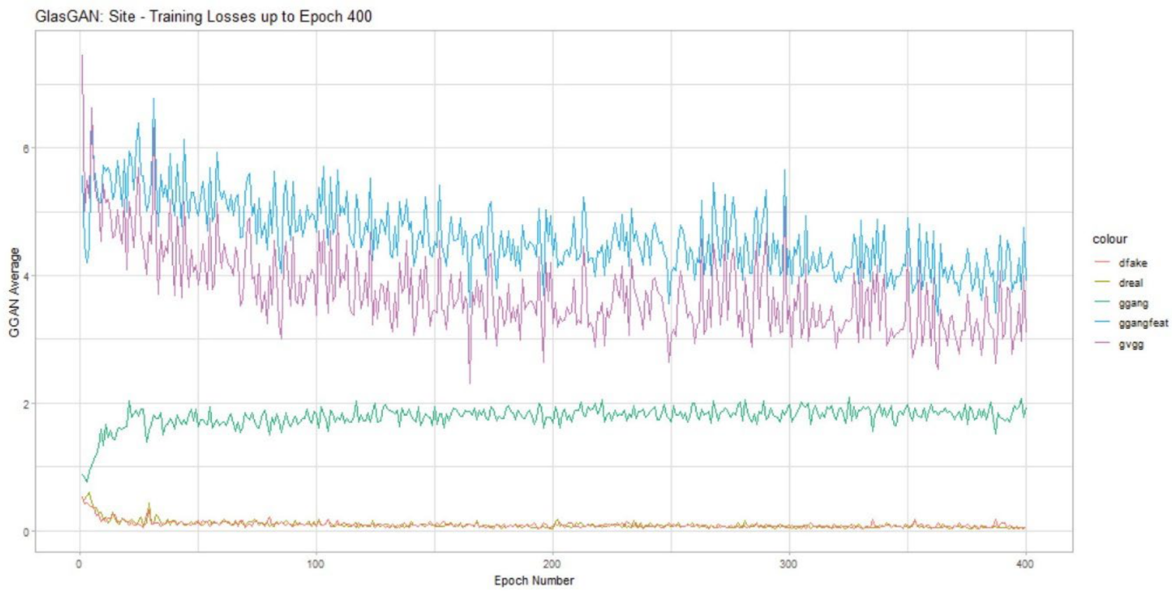


Figure 6 - Losses functions plot.

GlasGAN's training losses through epoch 400 showed DFake values quickly decreasing to around 0, only remaining fractionally above throughout training, see figure 6.0. A DFake value between 0.1-0.3 would be expected, with values close to 0 indicating the model is likely overfitted (Park and Zhu, 2020). The datasets created for the Site model were intentionally homogenous as means of reiterating vernacular architecture. Therefore, overfitting is not an issue for this model. The initial decrease, oscillation, and lack of extreme peaks with the GVGG and GGANGFEAT values indicate typical model training.

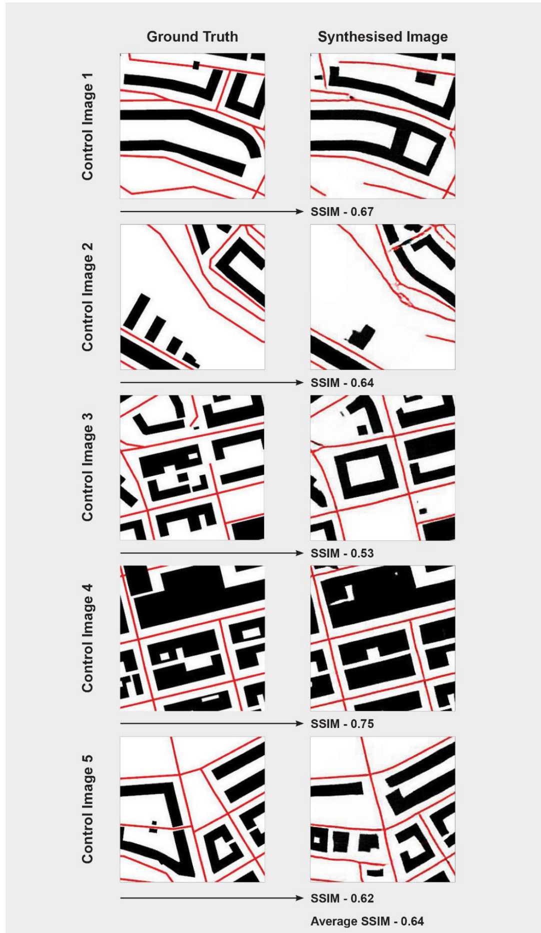


Figure 7 - Sample for SSIM Results.

The SSIM tests for the GlasGAN: Site control test images yielded values in the 0.53 - 0.75 range, with an average of 0.64 (Figure 7). This indicates a slightly positive degree of similarity, with outputs being more similar than dissimilar, illustrating successful model training.

3.3 Qualitative Validation

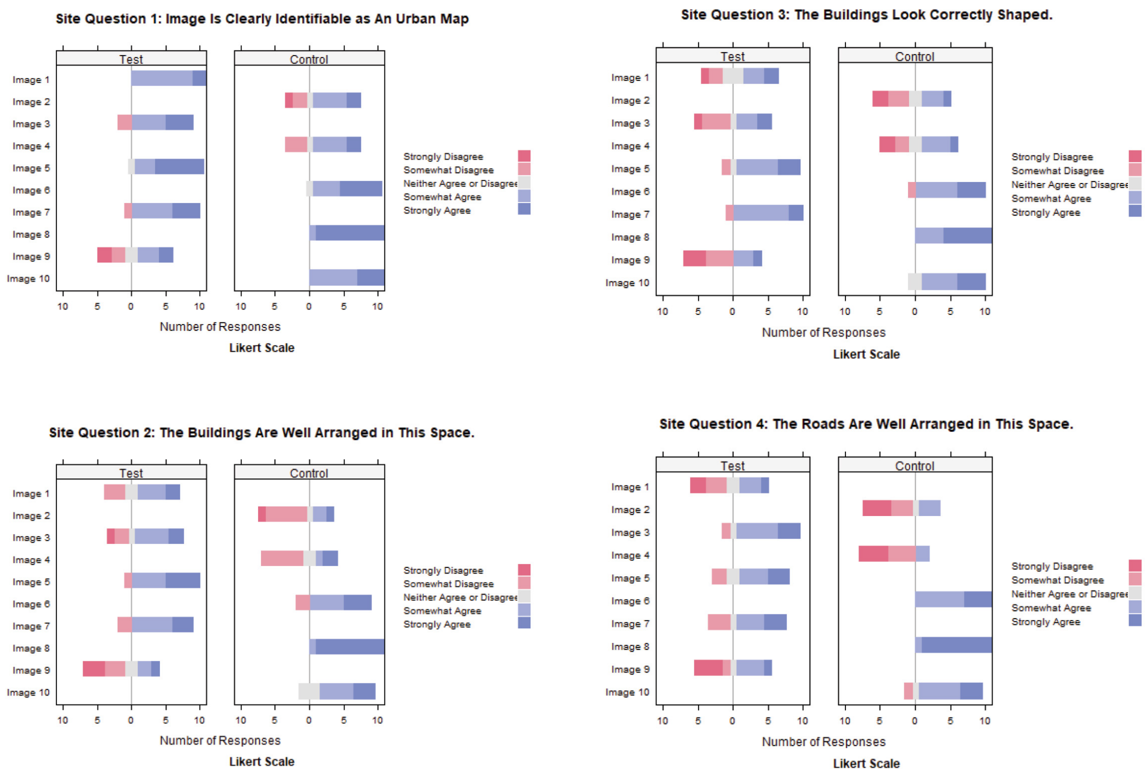


Figure 8 - Survey Results

The expert survey results in figure 8.0 showed similar responses for test and control images across all questions. All Shapiro-Wilk tests for individual answers data indicated a normal distribution (p-value > 0.05). Conversely, T-Tests comparing the average responses from a Likert scale between control and test images returned no statistically significant differences in survey respondents' perception of image quality between control and test outputs (p-values below 0.05 in most of the tests). This indicates the experts found the output quality in line with successful model training, supporting suitability for the urban design process.

4. Conclusions

The results of this study concluded that:

- The process was reproducible and was successfully expanded to include new data from Glasgow.
- The new complementary quantitative and expert-driven evaluations successfully validated the quality of the results.
- The new expert driven validation also supports the practical use of CGAN results by industry professionals.
- The additional methods are also open-source and reproducible.

References

Belam, C., Santos, L. and Leitao, A. (2019). *On the Impact of Machine Learning Architecture without Architects?* CAADFutures19.

Boim, A., Dortheimer, J. and Sprecher, A. (2022). A Machine-Learning Approach to Urban Design Interventions In Non-Planned Settlements. *CAADRIA proceedings*. doi:10.52842/conf.caadria.2022.1.223.

Khean, N., Fabbri, A. and Haeusler, H. (2018). Learning Machine Learning as an Architect, How to? In: *eCAADe 2018: Computing for a better tomorrow*.

Park, T., Liu, M.-Y., Wang, T.-C. and Zh, J.-Y. (2019). *Semantic Image Synthesis with Spatially-Adaptive Normalization*. NVIDIA.

Qu, Y., Chen, Y., Huang, J. and Xie, Y. (2019). Enhanced Pix2pix Dehazing Network. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. doi:10.1109/cvpr.2019.00835.

Raman, T.A., Kollar, J. and Penman, S. (2022). SASAKI: Filling the design gap—Urban impressions with AI. *Artificial Intelligence in Urban Planning and Design*, pp.339–362. doi:10.1016/b978-0-12-823941-4.00002-0.

Salehi, P. and Chalechale, A. (2020). Pix2Pix-based Stain-to-Stain Translation: A Solution for Robust Stain Normalization in Histopathology Images Analysis. In: *2020 International Conference on Machine Vision and Image Processing (MVIP)*. RAZI University.

Biographies

Alex Mackay: he is a design professional working in the fields of architecture and urban design. He recently completed an M.Sc in Urban Analytics at the University of Glasgow. His research focuses on the augmentation of architectural and urban design processes using machine learning methods.

Diego Pajarito Grajales: he is a Research Associate in Geospatial Data Science at the Urban Big Data Centre at the University of Glasgow working at the IDEAMAPS Ecosystem project funded by the Bill & Melinda Gates Foundation. His research currently relies on the design of geospatial technologies for data management with a citizen-centered approach.