Comparative Canvas
SPOT

Zaid, Inès; Nguyen, Ha Nhu; Kyprianou, Nikoletta; Karman, Tim

Publication date
2024

Document Version
Final published version

Citation (APA)
Zaid, I. (Author), Nguyen, H. N. (Author), Kyprianou, N. (Author), & Karman, T. (Author). (2024). Comparative Canvas: SPOT. Web publication/site, Medium.

Source: Nguyen, H. (2024, April 18). COMPARATIVE CANVAS : SPOT - Ha-Nhu Nguyen - Medium. *Medium*. https://medium.com/@h4nhu72/comparative-canvas-spot-620e967a2882

COMPARATIVE CANVAS : SPOT

by Inès Z., Ha Nhu N., Nikoletta K., Tim K.

Submission for:

AR2AA010 : AI in Architectural Design
at Delft University of Technology

Authors:

Inès Zaid
Ha Nhu Nguyen
Nikoletta Kyprianou
Tim Karman

Architecture Project:

The SPOT (Amsterdam, Netherlands)
Architect: KAAN Architecten

1. TL;DR

AI in architectural visualization is in its infancy; TU Delft's AI lab course teaches architects promising AI methodologies, an account.

2. MOTIVATION

As architects and architects-in-training, we gain spatial insight of the world surrounding us, as well as an understanding of the design process. Upon enrolling in the AI in Architectural Design course at TU Delft University, the nature of our case study, SPOT, sparked our analytical curiosity. Questions ensued: How does SPOT's design address diversity and conviviality? Could the design of a building contribute to such a goal? Would a scientific analysis of the building clarify such dynamics?

Integrating Artificial Intelligence (AI) into architectural visualization is an unexplored area. Existing AI models developed for comprehending and visualizing data broadly revolve around photo imagery for classification, object detection, or semantic segmentation tasks. Its use typically serves generic applications such as reverse image search, photo tagging, or prompt-based image generation (i.e.: Midjourney).

There is a significant disconnect between the rich database generated by architects (i.e.: floor plans and sections) and the capacity of current AI models to decipher this data to gain knowledge on the buildings at hand. A further gap lies in our own capacity (and lack thereof) as future architects to navigate these new methodologies. We were, therefore, enthusiastic about learning about AI-based computer vision and architectural visualization. The course promised to satisfy our curiosity about working with various data modalities. It offered strategies to navigate, manage, and interpret large urban datasets, which are often too complex for traditional analysis methods in an efficient fashion.

3. THE BUILDING

Fig. 1: The SPOT mixed-use complex. Source: KAAN Architecten

First things first: SPOT Amsterdam, designed by KAAN Architecten **(1)**, is an innovative mixed-use project in the Amstel III neighborhood, South-East Amsterdam. *(Figure 1)* The overarching goal of the project is to rejuvenate this office-centric district into a lively and diverse community with an urban densification strategy spearheaded by the integration of residential, work, and leisure spaces.

To start our investigation beyond the glossy website pictures and marketing information, we attended a presentation by Renata Gilio, managing director at KAAN Architecten. We learned that through its architectural qualities, SPOT has been designed to cultivate a vibrant urban atmosphere through a rich program integrating an offer of versatile areas featuring contemporary living needs with environmentally friendly practices. The architects fused tall architectural elements that played a double role: they placed the SPOT on the map like a landmark and a flagship of this district’s rejuvenation and provided its users with killer views of Amsterdam city skyline. The plan is configurated to optimize the accommodation of a broad range of activities, through a flexible layout that can be customized to meet tenants’ specific requirements, such as corporate offices or retail spaces.

SPOT is currently under construction, and our second basis for the investigation was to pay it a visit. We discovered the monochromatic nature of the area that this (yet to clad) building will address with its material pallet and dynamic shape. The building also incorporates methods to reduce the environmental impact of the building, through its considerable density, resulting in a green inner-city.

4. DEEP VISUAL ANALYSIS

Computer vision is an domain within the artificial intelligence field that teaches computers to analyze and process visual information. The machine can respond to visual observations it takes as an input digital imagery, such as those captured by cameras and temporal imagery such as videos. Deep learning models are used to transform these high dimensional inputs and to precisely identify and categorize objects and events.

Implementing computer vision algorithms for the SPOT, we were able to examine and visualize diverse datasets and tools. The basis of the methodology lies in the integration of aerial and street view images of the site and its direct surroundings, as well as detailed floor plan data from the building. To achieve this, we used tools,Programming language,libraries,modules, and deep learning models such as Python, Numpy, Pickle, Geopandas, and DINOv2 for deep-learning-based embeddings. Testing these tools we formed a GeoDataFrame that combined all data relevant sources. In turn, this allowed us to produce a large-scale plot of the building footprints specified by building identifiers on each footprint. Finally, the method assisted us in visualizing the spatial distribution and identification process of specific buildings. *Figure 2* depicts the large-scale plot and tile retrieval of 3 tiles, from the SPOT in the bottom right corner to the center of Amsterdam.

For our visualization project, we relied on a data-driven methodology to navigate and examine the large amounts of images, converting them into an easily understood format that allows for a deeper investigation of architectural characteristics. A data-driven methodology systematically extracts information and insights from organized (GeoDataFrames & JSON files) and unstructured data (Images & Graph data) , by applying algorithms, statistics, and machine-learning techniques.

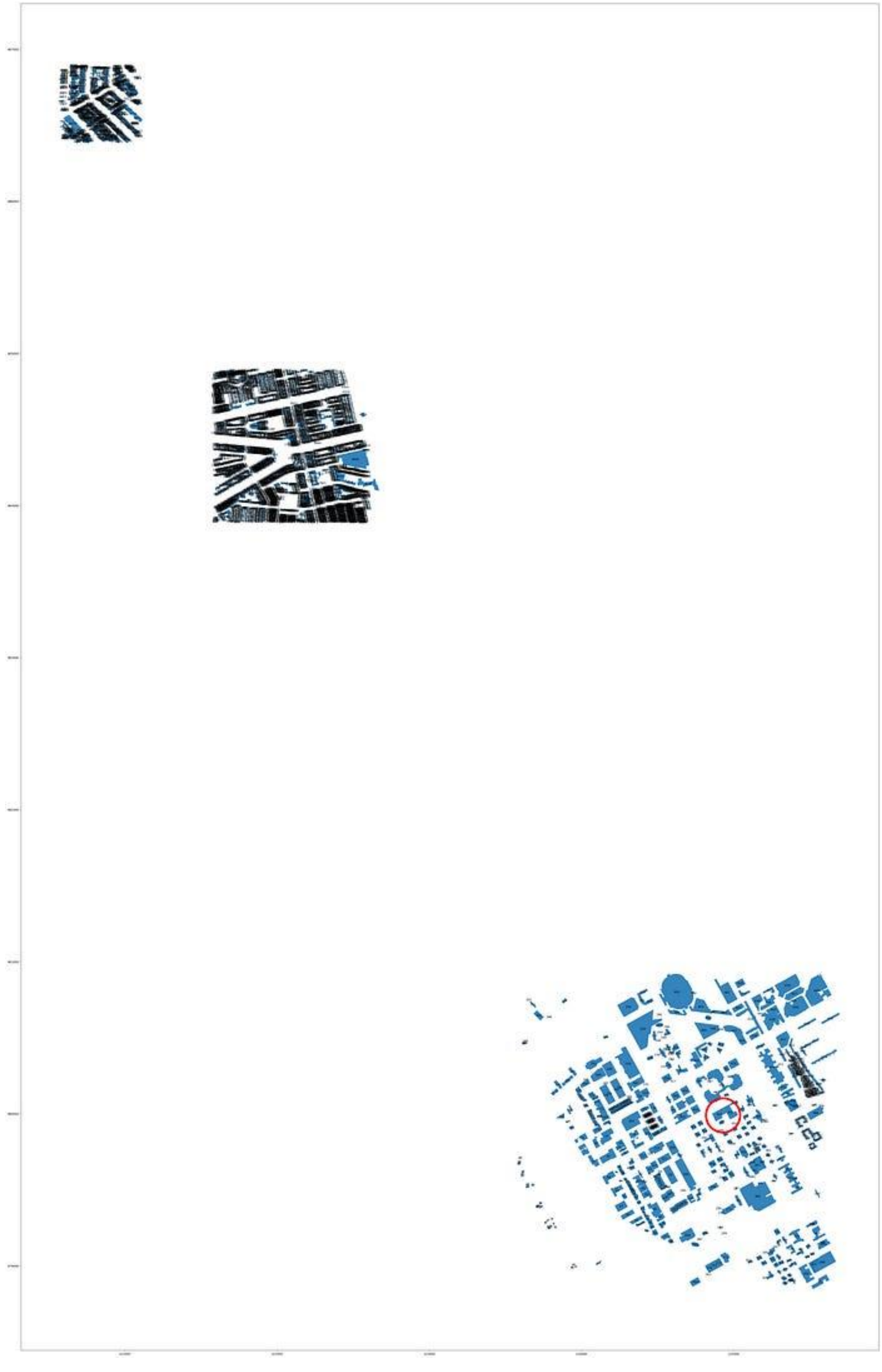

Fig. 2: Retrieval of three tiles from 3DBAG: Large scale Plot of three neighborhoods from Amstel II (SPOT circled in red) towards the center of Amsterdam

We generated and merged JSON files that map building indices to their geographical coordinates, leading to specific building selections, including images that may be missing or need updates. Embeddings must be defined as vector representations capturing each image's key characteristics. To develop the embeddings, we resized and transformed the images to fit into the DINO network model DINOv2 is an autonomous learning model that employs knowledge distillation to instruct transformers without the need for labeled inputs. A vector embedding is a numerical representation that transforms raw data into a vector space, where related data points are positioned in close distance to one another. We then concatenated the embeddings from both the aerial and the street view images and linked them to the vectors per image together. Thus, by using the DINOv2 model, the similarity of embeddings from each imagery gallery is obtained. For example, Figure 3 displays similarity analysis via DINOv2 of urban context using three different datasets: aerial, street view, and a combination of both titled as "urban similarity". This method of data processing using embeddings safeguards each track from data preparation to analysis, and then visualization is centered on precise and structured data input. This allows us to exhibit patterns and abnormalities in architectural styles and the expansion of cities as they evolve over time.

We then wanted to evaluate the large dataset obtained through Cyclomedia's API and satellite imagery, focusing on spatial configurations and functional effectiveness. To this end, we implemented Big Data analytics techniques. This refers to the analysis of large quantities of data using computational methods to identify patterns, trends, and connections, specifically in the field of human interactions and behaviors, here within the built environments. To effectively manage these extensive images and geospatial coordinates datasets, we utilized cloud storage, namely Google Drive, and employed advanced data processing libraries (Pandas and GeoPandas).

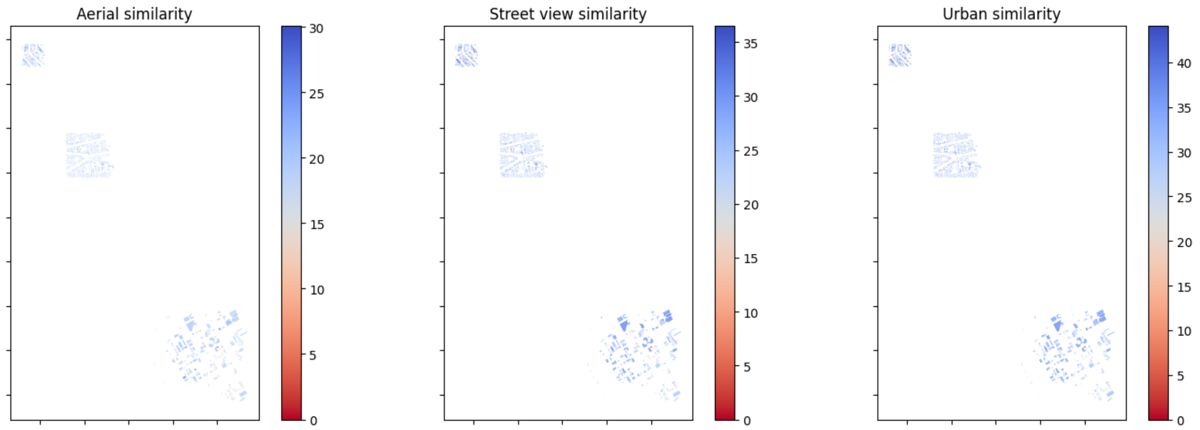

Fig.3: Visualization of urban context similarities on a 2D map from 3 different datasets. From left to right: aerial imageries, street view imageries, a customized dataset combining aerial and street-view imageries

The first step towards data-driven context analysis at neighborhood and urban scales involved transforming image arrays into vector embeddings using a pre-trained foundation model. Then, we used t-SNE (t-Distributed Stochastic Neighbor Embedding)[1]to further compress these embeddings by flattening them into more easily interpretable two-dimensional maps of data. Lastly , we visualized the image dataset on that 2D canvas on t-SNE coordinates by replacing points with the actual images. The resulting visualization of data points in the 2D map was more readable and, thus, allowed us to investigate the urban similarity of our query building and a collection of buildings in its surrounding neighborhood.

Incorporating computer vision, data-driven approaches, and big data into our architectural investigation gave us a detailed knowledge of SPOT's architectural core and its importance

in the urban context. That meant we could observe the project beyond the basic structural components, to analyze the complex interactions of color-coded visualization, combined data analysis on geospatial maps and highlight similarities in both aerial and street view context. The project, as seen on the poster, reveals how sophisticated analytical methods can convert unprocessed data into valuable insights that inspire novel building concepts and methodologies.

[1] t-SNE is a model used for dimensionality reduction to visualize high-dimensional data, particularly useful for datasets that cannot be easily represented.

5. TOOL & TECHNIQUES

Throughout this course, a variety of data modalities were used, including Python data formats, programming frameworks, AI models, and algorithms. The objective was to apply artificial intelligence to extract from these architectural data, the insights that were significant to us and to be able to display them in a visually compelling way. The primary data modality type used was graphical data depicting floor layouts. The data comprises nodes symbolizing rooms or areas within the floor plan, and edges that depict the connections between rooms, such as doors or open passages. These nodes are structured in graph formats. Therefore, demonstrating the capabilities of architectural illustration through graphs efficiently conveys intricate connections in simple forms to the outsider's eye. This method is akin to the adjacency diagram that architects use to communicate their early composition ideas.

Image data modalities showcasing a variety of perspectives, including aerial and street views were also used in this project. These images were subsequently converted into tensors, to facilitate the analysis using deep learning models. Additionally, we processed the urban layout geospatial data with GeoPandas. This process involved working with GeoPackage files containing vector data representing city buildings.

Throughout the project, Python data formats were used to store and manage the data. For instance, DataFrames and GeoDataFrames were used with Pandas and GeoPandas, allowing us to easily maneuver through structured data. Another example is the use of 'Tensors', specifically PyTorch tensors, which required extra caution when processing and managing image data, particularly during transformation and integration into neural network models.

We also employed graph data formats using the NetworkX library to analyze, manipulate, and generate complex networks from architectural floor plans. Moreover, we explored various Python modules and matrices essential for computation. One notable tool was 'Pickle,' which was used to load, store, and maintain complex data structures like graphs and tensors, effectively managing the model's input and output data.

Lastly, Google Colab was the main platform for its GPU capabilities, essential for deep learning applications, and facilitation for collaborative workflow. We employed PyTorch and torchvision to construct and deploy models, specifically for managing image data. Additionally, we applied UMAP (Uniform Manifold Approximation and Projection) to reduce the number of dimensions, allowing us to depict complex data efficiently. Collectively, these tools enhanced our processes for analyzing data and developing models.

For the Comparative Canvas, using Deep Neural Networks (DNNs) was important for extracting and analyzing image features, particularly in generating orientations and appearances. We also implemented Graph Convolutional Networks (GCNs) to examine graph-structured data from floor plans, identifying key characteristics of rooms (nodes) and their interconnections (edges). We employed dimensionality reduction techniques, especially t-SNE and UMAP, to visualize complex data. We chose the t-SNE algorithm for its ability to preserve local data structures, thus facilitating detailed visual comparisons. These techniques — DNNs, GCNs, t-SNE, and UMAP — are all integral components of our AI models and algorithms toolkit.

Lastly, for the Comparative Canvas, we visualized floor layout similarities using RPLAN dataset alongside the SPOT dataset for the building-level comparison. *(Figure 5)* The end result gave us some interesting insights into the layouts: residential units in SPOT were designed with a diverse range of typologies, evidenced by their layouts being positioned along the periphery and scattered all around the plot. In contrast, the RPLAN floor plans remain in the center of the plot, with unit layouts differentiating from one another in a gradual gradient.

To achieve our desired visualization, we tested out three iterations by increasing numbers of grids and simultaneously scaling up the plot everytime. In the first trial, we obtained a rather simplified plot, with many missing data points from RPLAN. As a result, the plot shows RPLAN unit layouts which were adjacent to one another not quite similar to their neighbours. *(Figure 5A)*. We then increased the number of grids and scaled up the plot, twice and then thrice *(Figure 5B and Figure 5C, respectively)* We observed that scaling up the grid brought more floor layouts closer in variations, which were previously lost in the more widely-spaced grid. The obtained visualized maps were thus enhanced with a better understanding of the datasets we had.

To this end, using UMAP on the data samples of SPOT only could have thrown a light ion the result. To reflect on the maps, the wide gap representing dissimilarities between SPOT unit layouts and RPLAN points at SPOT's diversity in units' typologies. Various balcony placements and their relationship to living room space, for example, are speculated to have contributed to this diverse range. To create a variety of residential units and to maximize the potential light reaching a unit, SPOT floor layout shows alternating units placed and rotated from one floor to the next, creating a varied positioning of balconies on the façade. This diversity is a meaningful architectural quality that has been visualized and can be interpreted from the map.

Fig. 5: Building Scale analysis: UMAP representation of SPOT and RPLAN dataset embeddings.

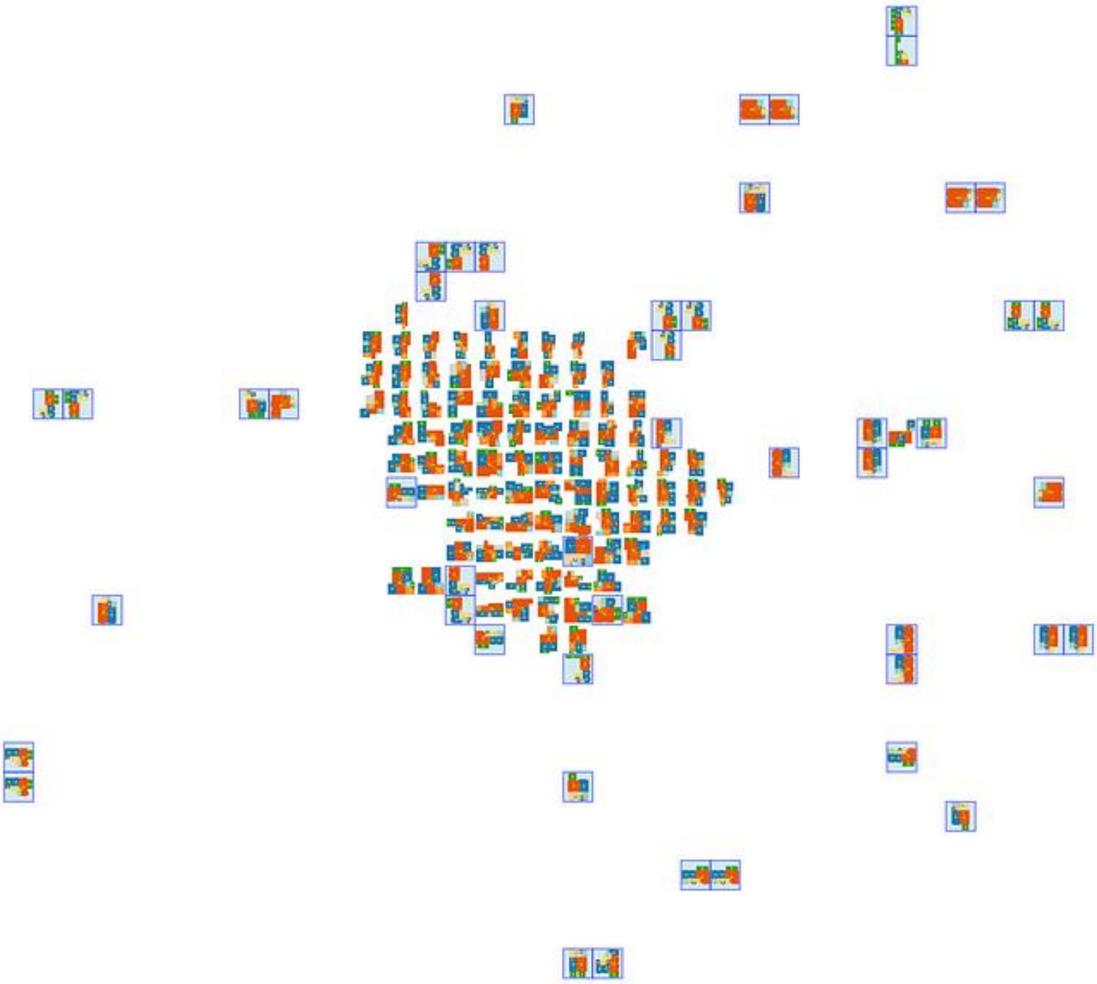

Fig. 5A: First trial using 40x40 grid

Fig. 5B: Second trial using 80x80 grid

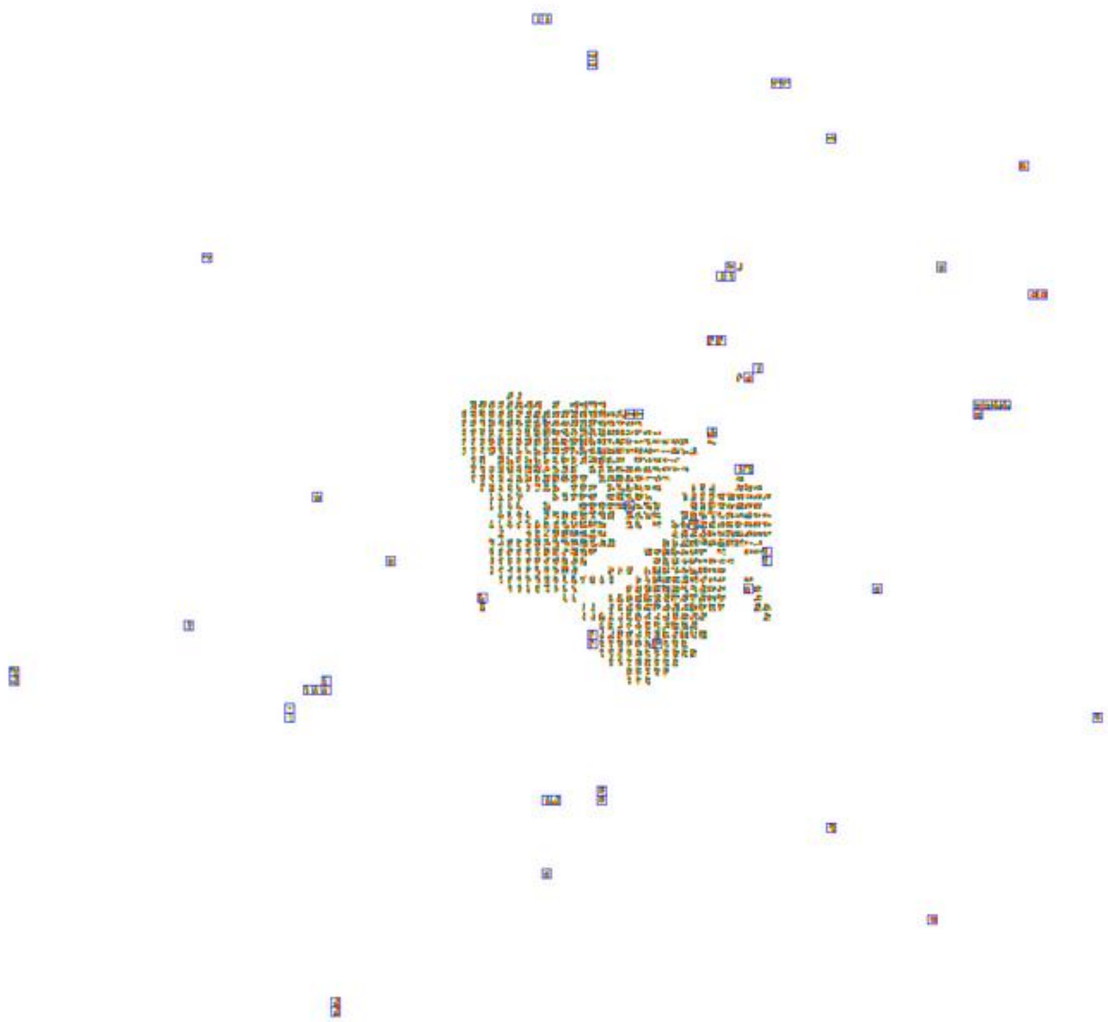

Fig. 5C: Final trial using 120x120 grid (shown in Comparative Canvas Poster)

6. REFLECTION

Using SPOT as a case study, the profound impact of artificial intelligence on visual analytics was explored through the implementation of AI models, using GCN and DNN. These tools supported the understanding of the complex dataset of a project unknown to us prior to the course, revealing its spatial connections within the units, the building, and the neighborhood. In fact, the graph reveals the logic of the building in a way that echoes the adjacency diagrams used by architects to communicate the choice of composition of a design.

A key takeaway from our work was learning to manipulate datasets to enhance visual outputs. Initially, our plots were simplistic, failing to accurately represent RPLAN data points and omitting some. By increasing grid numbers and scaling up plots, we captured more floor layouts that were previously overlooked in a widely spaced grid. However, scaling up reached a limit when it exhausted our RAM and hit a maximum resolution, underscoring the crucial balance between data resolution and system capabilities.

The application of GCNs appears promising in analyzing interconnected data from floor plans, as they can enhance our visual representation and comparison of complex compositions akin to the SPOT project. It is exciting to imagine how this can be applied to automated zoning evaluation or improving spatial efficiency in urban development projects. Furthermore, using DNNs to extract and compare characteristics from the site images has practical applications in fields like heritage preservation, where maintaining architectural characters is essential.

Our exploration revealed that various image inputs — like Cyclomedia exports, aerial images, and their combinations — yield significantly different results in urban analysis, highlighting how data types impact analytical outcomes. We also reflected on the transparency and ethics of using AI in visual analytics, as we used tools to support our re-ordering and presenting new data. The inherent opacity of the process poses challenges, as it creates a black box effect that makes understanding the analytical process difficult to assess, understand, and trust for reliability and ethical applications.

This leads to a final reflection: Just as diverse text prompts yield varied responses from ChatGPT, mastering AI tools and navigating different data types will enhance future architectural workflows, which are already evolving. Key questions arise: How can AI tools be made more user-friendly and navigable in architectural firms? How do we assess the performance of generative AI against its opaque nature? Lastly, what future possibilities does AI hold beyond image generation?

Overall, this work has given us a look into the keyhole to understand and apply AI logic to analyze and visualize complex architectural and urban data, on the one hand, and to reflect on the application and transparency of such technologies in the built environment.

References:

1. SPOT [Internet]. KAAN Architecten. [cited 2024 Apr 17]. Available from: https://kaanarchitecten.com/project/spot/