Evaluating the Geometric Properties of 2D City Layouts Using a Data-Driven, Deep Learning Approach

Sheung Chan (sc8635), Lazarus Chok (lcc9673), Yan Liang Tan (yt2217)

Word count (excluding references and appendices): xxx

Abstract

Reconstructing the geometric parameters of city layouts is complex and generally regarded as a qualitative task in many urban planning practices. Data-driven approaches are lacking as the current state-of-the-art cannot capture the hierarchical nature of 3D building data. To overcome these limitations, we use a proprietary tree-structured neural network ("AETree") developed by the AI for Civil Engineering Lab (AI4CE), which automatically reconstructs 3D building data. We hypothesize that the model's latent representations of spatial data are expressive of real-world city layouts. We will test this hypothesis by implementing AETree across New York City neighborhoods. We will also experiment with improving the model's scalability and using latent space interpolation techniques to create composite city layouts. Successful completion of the project will (1) represent a technical breakthrough for modelling spatial data, (2) provide urban planners with a practical tool to generate, compose and evaluate city layouts, and potentially (3) improve urban liveability and sustainability through data-driven urban planning.

Table of Contents

1. Introduction

By 2050, more than seven billion people will live in cities - nearly double today's urban population[1]. Rapid urbanization has led to hundreds of new cities being built globally, with many more being retrofitted, redeveloped, or vastly expanded. The current rate of urban land consumption will add nearly 500,000 square miles of urban area in the next 30 years[2]. Once an urban area is developed, its physical form would be locked in for generations.

The sheer magnitude of city building has increased the urgency for scalable urban design: well-planned cities that enable residents to develop a strong sense of place. Urban planning studies indicate that distinctive city layouts (the geometric parameters of urban spaces) are crucial for placemaking. However, it is inefficient and tedious for planners to design consistent city layouts as **they do not have a standard metric to quantify the geometric properties of city layouts**. Good urban design is difficult to scale.

Limited descriptors like street width, building heights, density do not consider the infinite geometric variety of city layouts. The difficulty of deriving a standard, objective metric (and therefore no common language) for describing city layouts can be traced to the sheer complexity of city layouts. Distinctive city layouts emerge from the relationships of urban elements such as paths, edges, districts, nodes, and landmarks[3]. Although there are metrics to quantify each element, there are no methods to evaluate city layouts comprehensively.

2. Literature Review

With unprecedented urbanization in recent decades, many cities are undergoing rapid redevelopment while urban expansion has been booming. The internationalization of planning solutions and mass manufacturing has resulted in conflicts between old and new developments, loss of sense of place[4], degradation of the quality of life[5], and growing environmental problems. The call for a more fundamental change and holistic approach to city-building has never been more urgent.

There is scope for positive change through the careful composition of urban layouts. Studies have shown that urban parameters such as density, compactness, and diversity of use correlate with the liveability and sustainability of the city[6]. Notably, maintaining continuity in the evolution of urban layouts could help sustain a sense of place and elevate quality of life[7]. Hence, the tools which urban planners have at their disposal to analyze urban patterns

---

[1] Ritchie, H. & Roser, M. (2018). *Urbanization*. Our World in Data. Retrieved from: https://ourworldindata.org/urbanization.
[2] The World Bank. (2020, Apr 20). *Urban Development*. Retrieved from: https://www.worldbank.org/en/topic/urbandevelopment/overview.
[3] Lynch, K. (1960). *The image of the city*. Boston: The MIT Press.
[4] Gokce, D., & Chen, F. (2016). 'Does the typological process help to build a sense of place?' *Urban Morphology*.
[5] Vijayakumar, Vinodh., Sangeetha, S.P. (2016). 'Urban form as a critical factor in the Quality of Life indicators – A review.' *In Materials Today: Proceedings.*
[6] Y.R. Jabareen. (2006). 'Sustainable Urban Forms: Their Typologies Models and Concepts' *Journal of Planning Education and Research.*
[7] Chen, F., & Thwaites, K. (2013). *Chinese Urban Design: A Typomorphological Approach*. Routledge.

become critical. They could empower planners to reliably reproduce desirable city forms that enshrine good urban planning principles and uncover connections with important local values, and potentially addressing larger liveability and sustainability issues.

There is strong theoretical basis for well-designed city layouts - urban planners including Jane Jacobs[8], Gordon Cullen[9] and Kevin Lynch[10] are strong proponents. However, these theories are not yet mainstream and are often cast as stylistic approaches. One reason for this underwhelming focus is that the current analytical process for classifying and evaluating urban typologies is "not entirely objective" and "lacks conceptual rigor".[11] Previous attempts to define a standard set of attributes to characterize urban layouts have been inconsistent and incomplete. Marshall developed a set of street typologies based on the configuration, composition and complexity of streets, but does not account for buildings and open spaces.[12] Meiner et al. proposed basic geometric attributes to describe building footprints, but focused on measuring urban density.[13] Recent attempts have expanded this set of attributes at the expense of increasing computational costs without addressing how different attributes may have different degrees of importance in urban outcomes.[14,15,16] While there is consensus within the planning community that urban layouts should be classified using numerical attributes, there is no agreement on what these attributes should be. Many either lack architectural precision (e.g., limited classification categories)[17] or are based on specific geographical regions and design traditions (e.g., UK design practice).[18]

Therefore, there is a need for a standard metric for city layouts that is generalizable across geographies and design traditions, and captures the complexity of city geometry without becoming too computationally expensive. Conceivably, recent advancements in deep neural networks could encode the geometric complexity of city layouts automatically and generate useful representations for urban planning applications. To the best of our knowledge, deep learning has not been applied to this problem because existing neural network architectures are not designed to understand complex spatial structures. Baseline evaluations (Section 4.3) show that state-of-the-art models like PointNet++ perform poorly in reconstructing spatial data.

This research addresses the long-standing problem of evaluating urban layouts objectively using a data-driven, deep learning approach for urban planners.

---

[8] Jacobs, J. (1993). *The death and life of great American cities.* Vintage Books.
[9] Cullen, G. (1971). *The concise townscape.* New York: Van Nostrand Reinhold Company.
[10] Lynch. Image of the City. *The MIT Press*.
[11] Gil, J., Beirão, J. N., Montenegro, N. & Duarte, J. P. (2011). 'On the discovery of urban typologies: Data mining the many dimensions of urban form', *Urban Morphology*, 16(1), pp. 27-40.
[12] Marshall, S. (2004). *Streets and patterns*. London: Routledge.
[13] Meinel, G., Hecht, R., Herold, H. & Schiller, G. (2008). 'Automatic derivation of basic urban structural data and integration in a geographic information system', Federal Office for Building and Regional Planning. Bonn, Germany.
[14] Gil. Discovery of Urban Typologies. *Urban Morphology*.
[15] Schirmer, P. M. & Axhausen, K. W. (2016). 'A multiscale classification of urban morphology', *Journal of Transport and Land Use*, 9(1), pp. 101-30.
[16] Bobkova, E., Pont, M. B. & Marcus, L. (2019). 'Towards analytical typologies of plot systems: Quantitative profile of five European cities', *City Science*, 48(4), pp. 604-20.
[17] Fialkowski, M. & Britner, A. (2008). 'Universal rules for fragmentation of land by humans', *Landscape Ecology*, 23, pp. 1013-22.
[18] Tarbatt, J. (2012). 'The plot: Designing diversity in the built environment', *A manual for architects and urban designers*. London: RIBA Publications.

3. Research Hypothesis

Our project's central hypothesis is that complex 2D city geometry can be objectively encoded using deep learning techniques that preserve the city's spatial hierarchy. We also hypothesize that these encoded features can become faithful representations of the complex geometric features and are useful for urban planning applications. The research questions and hypotheses are elaborated in **Table X**.

Technical stage	Problem statement	Research question	Hypothesis
Feature Engineering			
Encoding city layouts using AETree[19]	2D city layouts contain complex geometry that cannot be easily reduced to simple, standard metrics for efficient urban planning	How can we develop a standard metric to quantify the geometric properties of city layouts?	AETree can encode the complex geometry of city layouts, such that their latent features are statistically representative of their basic geometry
Urban Planning Applications			
Classification of Urban Typologies	There is no standardised, data-driven way of classifying existing and new urban typologies, making it inefficient for urban planners to objectively describe city layout designs	How can we objectively classify urban typologies based on their complex geometries?	Clustering the latent features of city layouts is an objective way of classifying urban typologies based on their complex geometry
Measuring neighborhood layout geometries	There is no standard, data-driven metric for measuring geometric properties of city layouts, making it impossible to quantitatively compare neighborhood layouts	How can we develop a standard metric for quantifying the geometric properties of city layouts?	The diversity of city layouts in each neighborhood can be represented by a single component
Blending city layouts	There is no systematic and reproducible way of blending city layouts to create composite, conceptual layouts	How can we generate composite urban patterns that blend distinct city layouts?	Interpolated latent vectors of two latent spaces can be decoded to generate a composite city layout

Table x. Research questions and hypotheses.

3.1 Project Overview

[19] AETree is AI4CE's proprietary deep neural network for spatial data.

To develop a standard metric for quantifying city layout geometry targeted at urban planners, our project will (**Figure x**):

1. **Data preparation and model engineering.** Encode geometric features of a geo-referenced dataset of Manhattan city layouts using the AETree neural network,
2. **Feature engineering.** Engineer features that faithfully represent city layout geometries using the encoding,
3. **Urban planning applications.** Develop three proof-of-concept urban planning applications using these features.

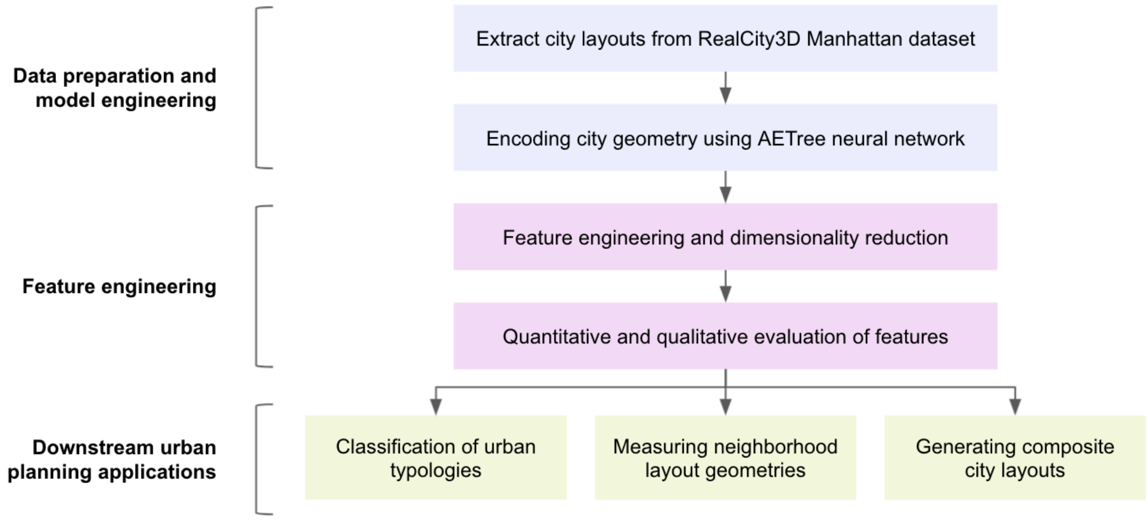


Figure x. Overview of project methodology.

4. Data Preparation and Model Engineering

4.1 RealCity3D Dataset

RealCity3D[20] is a large-scale georeferenced 3D shape dataset of real-world city buildings in NYC. We developed benchmarks for RealCity3D, including using the dataset for city layout generation. We submitted our findings to the 35th Conference on Neural Information Processing Systems.

To derive 2D city layouts from the RealCity3D dataset, each building polygon is transformed into a minimum 2D bounding box. 45,000 geo-referenced buildings in Manhattan were batched into distinct city layouts comprising of 32 neighboring buildings.

[20] Wen, C., Han, W., Chok, L., Tan, Y. L., Chan, S. L., Zhao, H. & Feng, C. (2021). 'RealCity3D: A large-scale georeferenced 3D shape dataset of real-world cities.' Retrieved from: https://ai4ce.github.io/RealCity3D/.

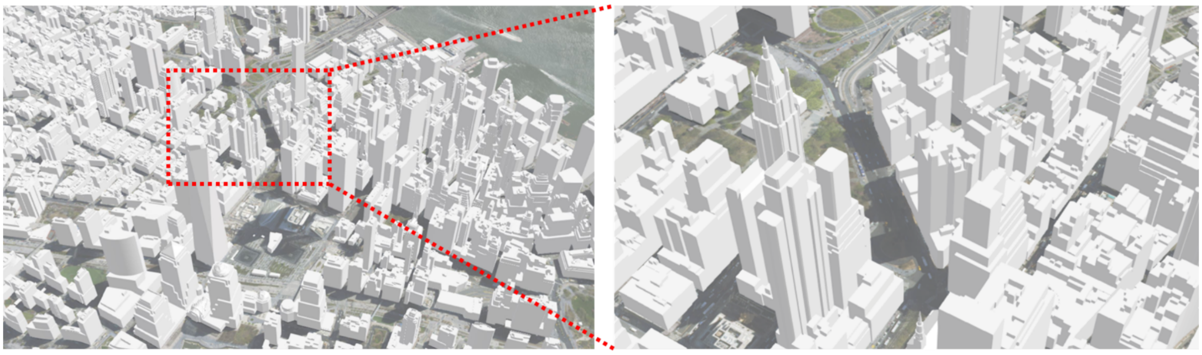


Figure 1. Detailed views of 3D building shapes in New York City.[21]

4.2 AETree Neural Network

AI4CE's proprietary tree-structured neural network AETree[22] is used to encode the complex 2D city layout geometry into a shared latent space, from which standard features can be engineered for urban planning applications. Training the AETree autoencoder requires tuning of hyperparameters, such as the learning rate and loss functions. The autoencoding process has three stages:

Data processing and hierarchical clustering. Our dataset contains 45,000 Manhattan city layouts, where each building footprint in the layout is represented by a tensor of (x, y, l, w, a), where x and y denote the center coordinates of the footprint, and l, w, a denote the length, width and orientation angle of the footprint respectively. AETree's hierarchical clustering algorithm transforms the spatial data into a tree-structured array of vectors, which preserves spatial information in a hierarchical manner. This process combines the boxes which are closest to each other until a final cluster emerges - representing the features of the entire 32-box plot (Figure 2). 70% of the dataset are taken as training sets, 10% as validation sets, and 20% as test sets.

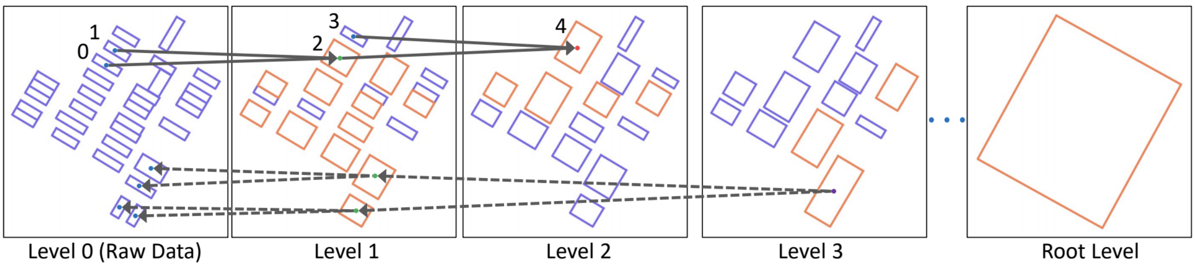


Figure 2. Tree-structured data.[23]

Encoder. The encoding stage uses the LSTMCell architecture as the encoding function f_e to learn the long-range dependency of features at each layer of the binary tree, from bottom to top. Each node in the tree is represented by its geometric parameters (i.e., relative centroid coordinates) P and a feature representation F. Given the parameters of the left and right children nodes, the parent node's feature $F_p^{'}$ can be computed as:

---

[21] Ibid. Retrieved from: https://ai4ce.github.io/RealCity3D/.
[22] Feng, C. (2021). 'Urban Areal Spatial Data Generation Using Auto-Encoding Tree'. Unpublished.
[23] Ibid. Unpublished.

$$F'_p = f_e([P_l, F_l]) + f_e([P_r, F_r])$$

By concatenating all encoding features to obtain the root node feature F_{root}, AETree creates a latent representation that preserves the spatial structure of the city layout.

Decoder. The decoding stage hierarchically decodes the root node feature to reconstruct the parameters of the original city layout (Figure 3). At each level, the parent node is decoded into two children nodes:

$$[P'_l, F'_l, L'_l, P'_r, F'_r, L'_r] = f_d([P'_p, F'_p])$$

Where f_d is the decoding function and P', F', L' are a node's decoded geometric parameters, feature representation and indicator of being leaf nodes or not. When all nodes are identified as leaf nodes, decoding will stop. L1 and binary cross-entropy are used as loss functions during the training phase.

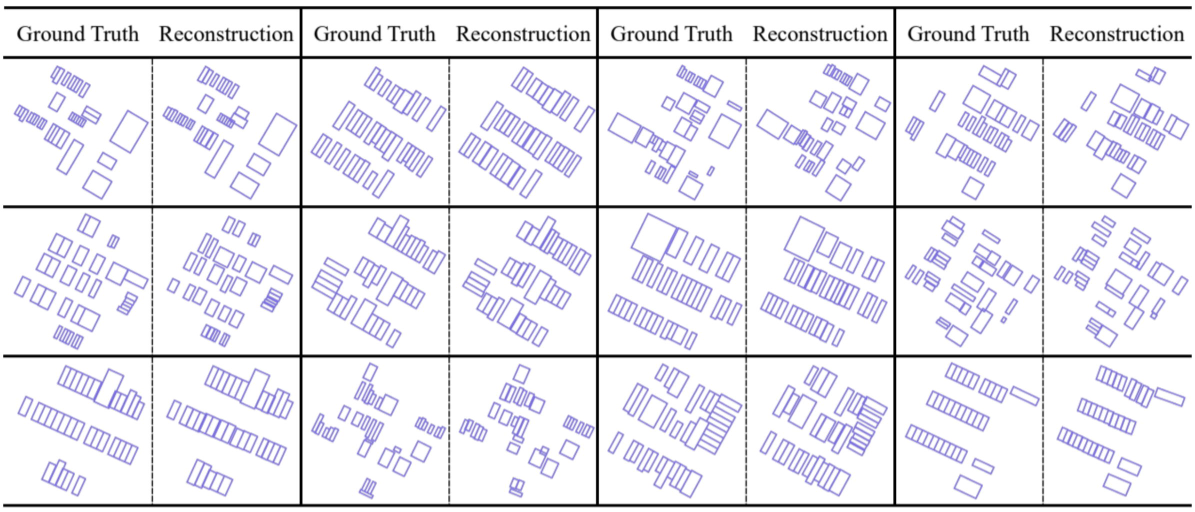


Figure 3. Example of AETree before and after autoencoding.[24]

4.3 Benchmarking AETree performance

We benchmarked AETree's city layout reconstruction against two other popular generative methods: a procedural modelling approach (ESRI CityEngine) and a state-of-the-art data-driven approach (PointNet++).

4.3.1 CityEngine

ESRI CityEngine is a commercial software that uses a procedural modeling approach based on L-systems to create large-scale city models. By creating road networks and dividing the

[24] Ibid. Unpublished.

parcels into lots, it generates buildings on the allotments using predefined rules. The building footprint is generated using default rules with extensive manual adjustment of parameters.

4.3.2 PointNet++

PointNet++[25] is a deep learning algorithm that captures local structures hierarchically and preserves the 3D spatial features of point clouds. By transforming the corners of each building footprint into points, we were able to apply PointNet++ to recursively partition and group the point data. The encoding stage recursively partitions and groups the input data. We then use the set segmentation approach along with Multilayer Perceptron (MLP) to decode the grouped data points and reconstruct the 2D city layout.

4.3.3 Benchmarking results

The baseline methods do not perform well in comparison with the AETree model. To intuitively show model performance, we randomly select some generation results of each model, as shown in Figure 4.

Generated layouts from the AEtree model are superior compared to PointNet++. The results from PointNet++ show irregular shapes of building footprints far less refined than the results generated using the AETree algorithm. While CityEngine is able to generate well-ordered city layouts, it loses style variance due to its rigid set of predetermined parameters and rules.

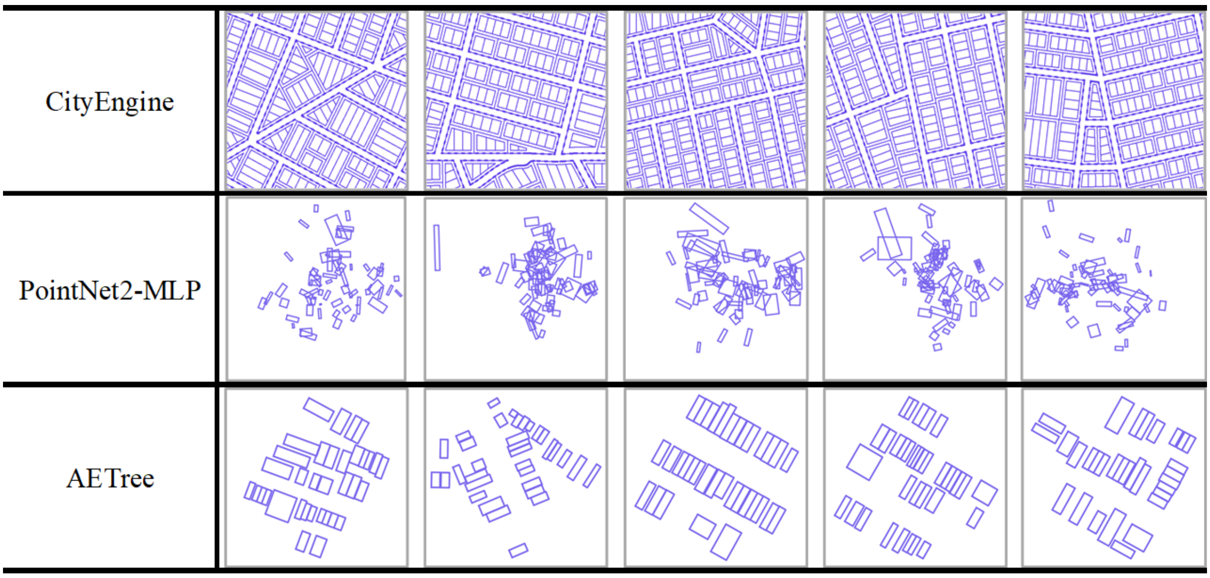


Figure 4: City layout generation results of the models trained on NYC dataset.[26]

4.4 Limitations of the AETree model

Although AETree successfully reconstructs 2D city layouts and is superior to state-of-the-art procedural generation and other data-driven methods, there remain significant limitations:

---

[25] Qi, C., Yi, L., Su, H. & Guibas, L. J. (2017). ‘PointNet++: Deep hierarchical feature learning on point sets in a metric space’, *Conference on Neural Information Processing Systems 2017*.
[26] Feng, Urban Areal Spatial Data Generation. Unpublished.

Limitation	Proposed solution[27]
AETree can only reconstruct a limited scale of 32 buildings. This limits the scalability of the model to reconstruct larger urban areas.	1. Using the existing LSTMCell architecture, adjust hyperparameters, e.g. loss function, to increase the output to 64-building layouts. 2. Replace the LSTMCell with an attention-based transformer architecture, which is more efficient in learning long-range dependencies for sequence-to-sequence tasks. Results are included in **Appendix X**.
Building footprints are represented by primitive 2D bounding boxes. This ignores roofscapes, facades and setbacks, which are important for urban patterns.	This might be overcome by enlarging the feature representation F to include more building features.
The model is only trained on Manhattan building data and may have limited applicability to cities with different city layouts.	We can create domain-invariant representations of city layouts by including diverse cities in the training set.

5. Feature Engineering

After preparing a geo-referenced dataset of 45,000 2D city layouts in Manhattan and training the AETree model on the Manhattan dataset, we extracted 512 latent features for each of the 45,000 city layouts. The latent features are taken from the last hidden state vector in the LSTM network, where it is assumed that all useful spatial information has been encoded.

However, a tensor of 512 features is too bulky for downstream applications. We applied dimensionality reduction using principal component analysis (PCA) and evaluated - both quantitatively and qualitatively - whether the reduced latent space correlates with basic geometric features of city layouts (**Figure x**).

[27] Limitations 2 and 3 are beyond the scope of this project.

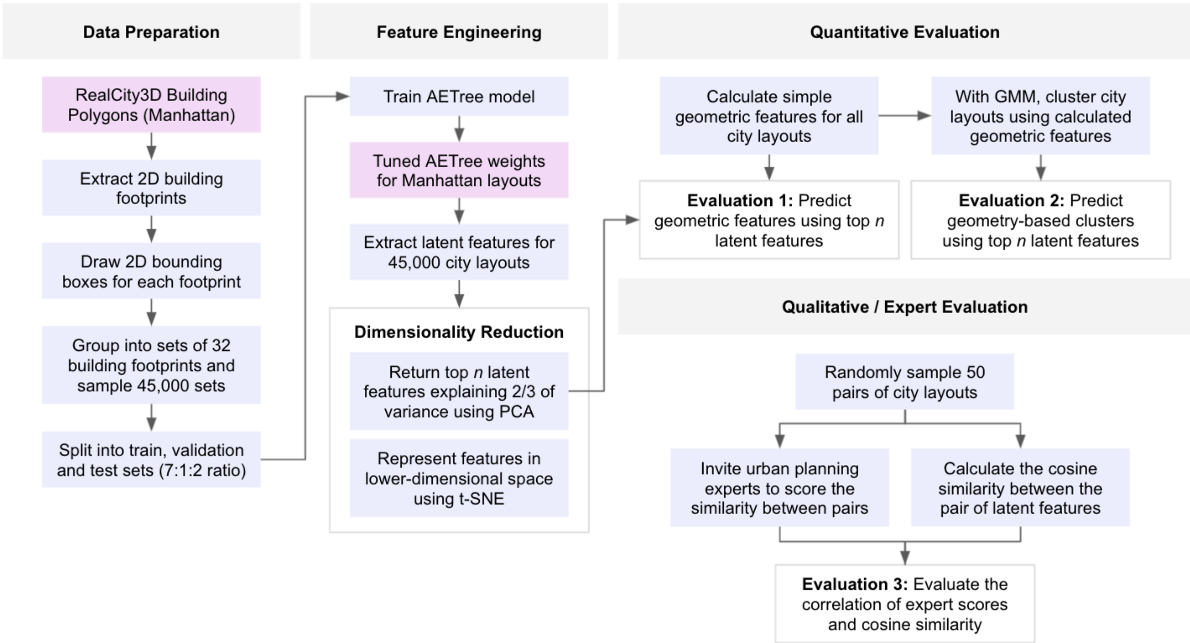


Figure X: Overview of feature engineering and evaluation.

5.1 Dimensionality Reduction of Latent Space

PCA is a feature extraction technique that maps the original latent space onto a lower-dimensional feature subspace. 512 latent features are compressed to 50 principal components, explaining 68.8% of the dataset variance.

For visualizing the feature subspace, we used t-distributed stochastic neighbor embedding[28] (t-SNE) which preserves local clustering structures. Our t-SNE plots were used to visualize distinct urban typologies in urban planning applications (Section 6.1), similar to Rhee et al.'s urban analysis of Pittsburgh.[29]

5.2 Quantitative Evaluation of Latent Features

To validate that the latent features are robust representations of the urban patterns, two qualitative tests were carried out using the GMM cluster labels:

Evaluation 1: Predict geometric features using top n latent features

The first qualitative test uses four machine learning (ML) methods to predict values of individual geometric features. Zero Rate and linear regression were used as accuracy baselines to evaluate the performance of the ML models. Four metrics, R Square, Mean Square Error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), were reported for the baselines and ML models. Based on the results for perimeter length in Figure X, the ML models performed significantly better than the baselines. Similar results

[28] van der Maaten, L. & Hinton, G. (2008). 'Visualizing data using T-SNE', *Journal of Machine Learning Research*, 9.
[29] Rhee, J., Llach, D. C. & Krishnamurti, R. (2019). 'Context-rich urban analysis using machine learning: A case study in Pittsburgh, PA', *Carnegie Mellon University*.

were achieved for average, standard deviation and range of building length/width ratio, indicating that the compressed latent features were able to predict individual geometric features with reasonable accuracy.

Geometric Feature – Perimeter Length					
s/n	ML Method	R2 Score	MSE*	RSME*	MAE*
	Baseline 1 - ZeroR	0.000	2.252	1.500	1.138
	Baseline 2 – Linear Regression	0.012	110.679	10.520	10.481
1.	Decision Tree	0.310	1.578	1.256	0.940
2.	Random Forest	0.499	1.011	1.006	0.772
3.	SVM	0.694	0.700	0.837	0.658
4.	KNN	0.439	1.282	1.132	0.865

* The lower the score, the better the fit of the model

Figure X. Model performance for perimeter length prediction using latent space

Evaluation 2: Predict geometry-based clusters using top n latent features

To facilitate the second evaluation, we used Gaussian mixture model (GMM) to perform clustering on 17 geometric features of the urban patterns. Jensen-Shannon (JS) metric[30] and Bayesian Information Criterion (BIC)[31] were used as the cluster performance evaluations to determine the optimal number of clusters, which was determined to be ten clusters.

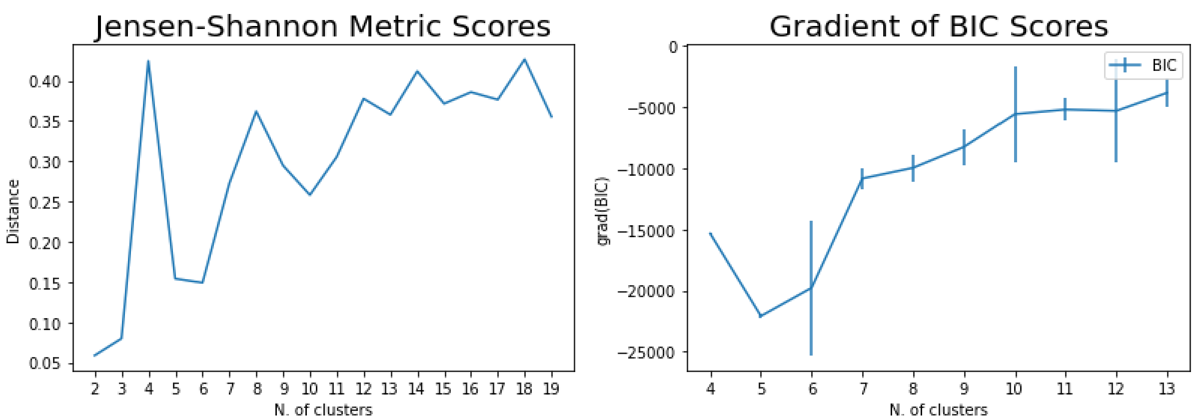


Figure X. GMM cluster performance evaluation using JS and BIC metrics

We used five different ML methods to predict the classification of latent features, and the prediction accuracy is reported based on geometric feature cluster labels. Zero Rate and

[30] Endres, D. M. & Schinden J.E. (Jul 2003). 'A New Metric for Probability Distributions', *IEEE Transactions on Information Theory, Vol. 49, No. 7*
[31] Schwarz, Gideon E. (1978). 'Estimating the dimension of a model', *Annals of Statistics, 6 (2): 461–464*

Random Rate Classifiers were used as accuracy baselines[32] to evaluate the performance of the ML models. Based on the results, all the ML models performed better than the baselines, thus validating that the compressed latent features can be used to explain individual geometric features.

s/n	ML Method	Best Score	Baselines	
			Random	ZeroR
1.	Logistic Regression	0.503	0.107	0.356
2.	Decision Tree	0.399		
3.	Random Forest	0.461		
4.	SVM	0.488		
5.	KNN	0.448		

Figure X. Model Accuracy of Cluster Prediction using Latent Space

5.3 Qualitative / Expert Evaluation of Latent Features

In addition to confirming that the reduced latent space is statistically representative of basic geometric features, we sought expert validation on the *statistical latent distance* between city layouts. If we assume that AETree can encode geometric complexity, then the statistical distance between two city layouts should correspond with an urban planner's visual interpretation of their similarity. Achieving expert validation would further prove the real-world relevance of our approach.

50 pairs of city layouts were randomly selected for expert validation. We invited eight professional urban planners to score the similarity between city layouts on Qualtrics. The cosine similarity, which measures the cosine of the angle between two vectors projected in a multi-dimensional space, was calculated for the 50 pairs by taking their reduced latent space as input.

Evaluation 3: Evaluate the correlation of expert scores and cosine similarity

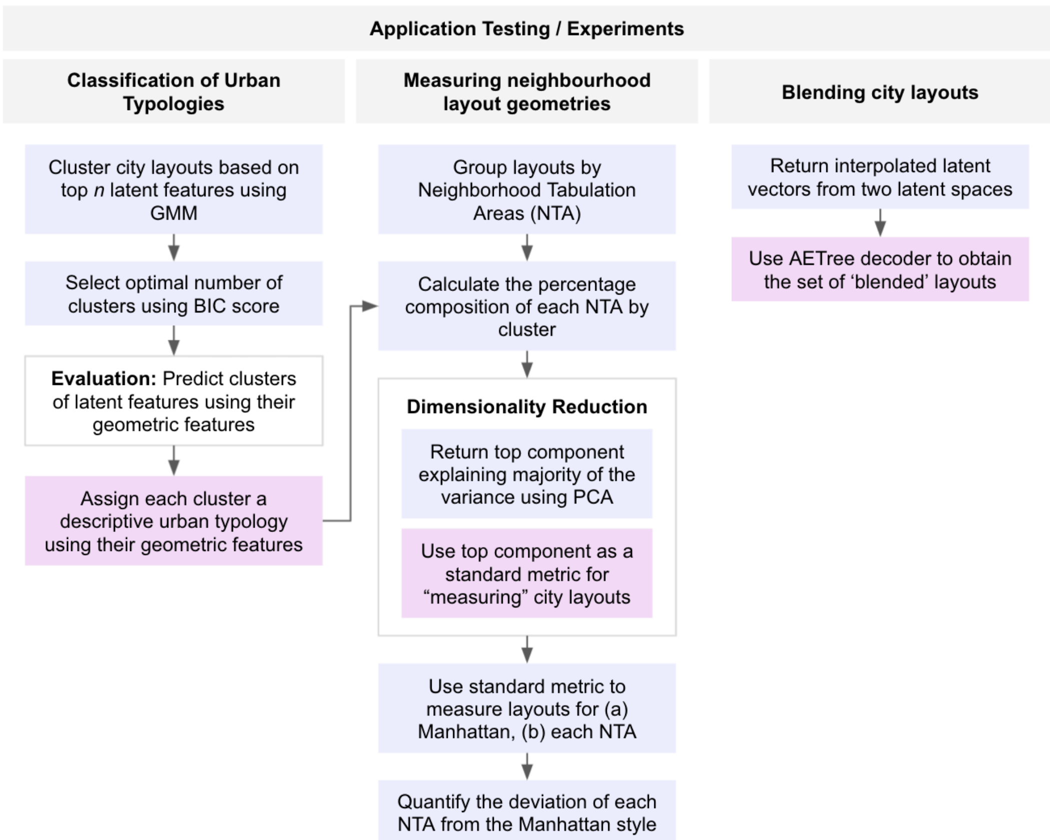
We assume that individual urban planners may have a subjective frame but their scores will follow a similar distribution, therefore we standardized their scores. There was a 43.1% correlation between the average of standardized scores and the cosine similarity of pairs (**Figure X**), suggesting a moderate positive correlation between expert judgement and the statistical latent distance between city layouts.

[32] Towards Data Science. (May 2021). 'Choosing a Baseline Accuracy for a Classification Model', Retrieved from *https://towardsdatascience.com/calculating-a-baseline-accuracy-for-a-classification-model-a4b342ceb88f*

Figure X. Correlation of expert scores and cosine similarity.

Taking the mean scores as threshold, we simplified the correlation to a binary test (similar vs. not similar). There was consensus between expert scores and statistical distance 66% of the time, which is better than random. This, together with our quantitative evaluation, justifies using the reduced latent space for urban planning applications.

6. Urban Planning Applications



6.1 Classification of Urban Typologies

6.1.1 Clustering of Latent Space using using Guassian Mixture Model

Similar to the evaluation done to predict geometry-based clusters using top n latent features, we used GMM to perform clustering on the 50 principal components for latent features. JS metric and BIC were used as the cluster performance evaluations and identified eleven clusters to be the optimal number of clusters .

6.1.2 Latent Space Cluster Prediction using Geometric Features

We used five different ML methods to predict the classification of geometric features, and the prediction accuracy is reported based on latent space cluster labels. Zero Rate and Random Rate Classifiers were used as accuracy baselines to evaluate the performance of the ML models. While the results show that all the ML models performed better than the baselines, the performances were only moderately better than Zero Rate Classifier. This is in contrast to the significantly better model predictions for geometry-based clusters using latent features. We surmise that this difference could be due to the inherent nature of latent features encapsulating more information on urban patterns.

s/n	ML Method	Best Score	Baselines	
			Random	ZeroR
1.	Logistic Regression	0.196	0.0967	0.154
2.	Decision Tree	0.164		
3.	Random Forest	0.186		
4.	SVM	0.203		
5.	KNN	0.159		

Figure X. Classification model accuracy for cluster prediction using geometric features

6.1.3 Feature Contributions using Shapley Values

An interpretability method, Shapley Additive exPlanations (SHAP)[33], was used to determine the contribution of each geometric feature and explain the output of the machine learning models. The explanation model is based on an additive feature attribution method to estimate the Shapley value of each geometric feature. We compute the Shapley values for the five ML methods used to predict latent space clusters.

Based on the Shapley values for the best performing model (SVM), the total building footprint perimeter length, average and standard deviation of building footprint length / width ratio are found to be the highest contributing features for ML predictions. The top five contributing features are also consistent for the other four ML methods, detailed in Appendix XXX.

[33] Scott Lundberg. (2018). SHAP (SHapley Additive exPlanations). Retrieved from https://shap.readthedocs.io/en/latest/index.html

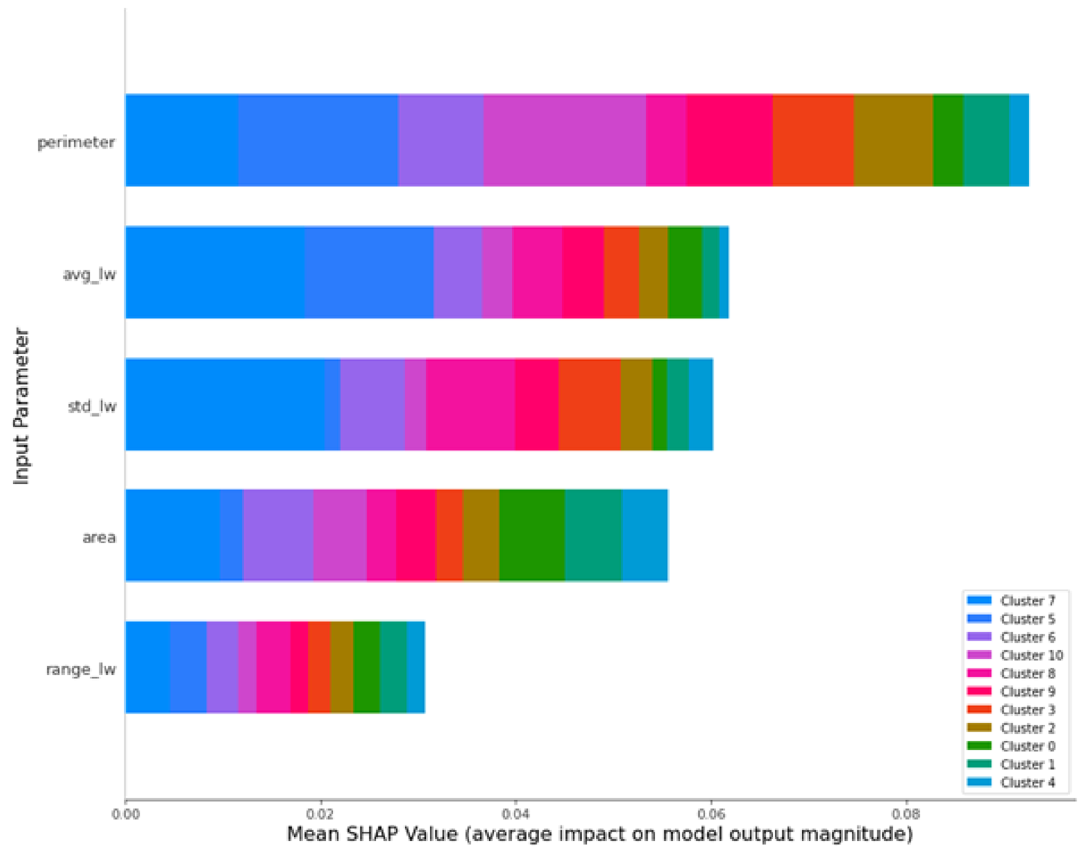


Figure X. SHAP values for the top contributing geometric features in the SVM model.

6.1.4 Statistics by Latent Space Clusters

Based on the latent space clusters that were identified using GMM, we further investigated if we were able to identify distinct clusters using the summary statistics of the top contributing geometric features, detailed in Table X. It is striking that some clusters have uniquely high and low values for some geometric features, which indicates that the clusters were not random and could be classified based on their geometric properties. For example, we would expect a cluster with a high total footprint area to fit the profile of higher density urban patterns.

Cluster	Building Area	Perimeter Length	Average Length/Width Ratio	Standard Deviation of Area	Standard Deviation of Perimeter Length
1	-0.203	-0.295	0.369	0.554	0.563
2	-0.383	-0.108	0.394	-1.052	-0.471
3	-0.124	-0.044	0.152	-0.256	-0.093
4	-0.601	-0.512	0.231	-0.750	-0.084
5	0.513	0.391	-0.343	-0.131	-0.256
6	-1.220	-1.204	0.451	0.715	0.962
7	0.562	0.174	-0.352	1.175	0.613
8	0.250	-0.022	0.140	0.866	0.567
9	-0.497	-0.271	0.947	-0.160	0.012
10	2.526	2.698	-2.768	-2.032	-2.655
11	-0.823	-0.806	0.781	1.071	0.842

Figure X. Standardised mean values of geometric features for each cluster.

To validate this claim, for each geometric feature of interest, we sampled several urban patterns with scores close to the overall average score of the cluster and visually compared them with samples from another cluster, as presented in Figure X. The visual results intuitively validate the expected profiles of the cluster to a reasonable degree:

Area Density - Cluster 6, which has a lower average building footprint area and length of building perimeter, has smaller buildings with larger spacings between them, occupying less space within the frame's limits. Cluster 10, which has a higher average building footprint area and length of the building perimeter, has larger building footprints and tighter spacings between them.

Building Proportion Regularity - Cluster 9, which has an average building length/width ratio that deviates the most from 1, tends to have more stretched and elongated buildings. Cluster 7, which has an average building length/width ratio close to 1, has more buildings that are more squarish and regular.

Pattern Evenness - Cluster 8, which has a higher standard deviation for building area, tends to feature a wider variety of building sizes in its urban patterns. Cluster 4, which has a lower standard deviation for building area, has more similarly sized buildings.

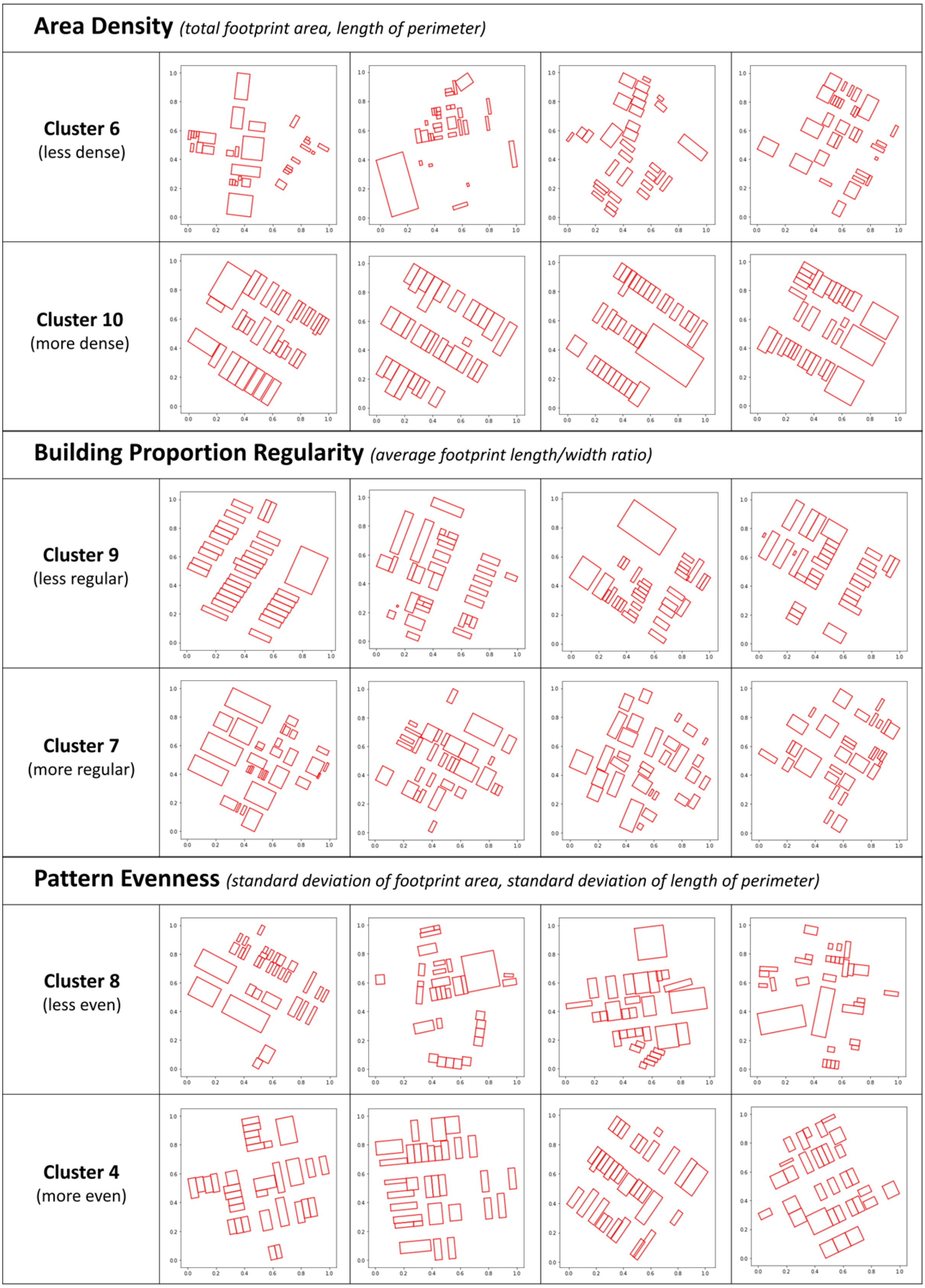


Figure X. Profiles of urban pattern clusters in terms of geometric features.

6.2 Measuring neighborhood layout geometries

Having identified unique urban typologies in Manhattan, our approach allows urban planners to further characterize a neighborhood's typological signature. During redevelopment of existing neighborhoods, the computational analysis of neighborhoods allows planners to objectively assess how to retain the neighborhood's unique urban character. Proposals for new city layouts can be evaluated against how they change the neighborhood's typological signature, in relation to surrounding neighborhoods and the city's overall urban profile.

6.2.1 Calculating Manhattan's Typological Signature

Establishing a baseline urban profile for the whole of Manhattan contextualizes subsequent neighborhood analysis. Using the optimized GMM clustering of city layouts (Section 6.1), the city's (or neighborhood's) typological signature is calculated by its percentage composition by cluster. Manhattan has a high percentage of cluster 8 and 10 layouts (**Figure x**, see detailed cluster descriptions in Appendix), which are both characterized by moderate density, even grain, and consistent orientation - emblematic of the classic Manhattan grid. Conversely, sparse layouts (e.g., clusters 0, 2, 6, and 9) are not as well-represented.

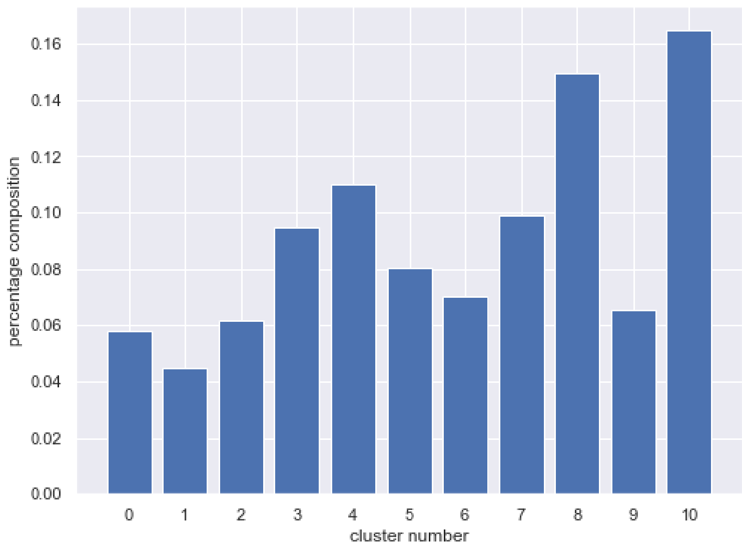


Figure x. Manhattan's urban typological signature.

6.2.2 Measuring neighborhood typological signature and deviation from baseline

Similarly, the typological signature can be calculated at the neighborhood level. The 45,000 city layouts were grouped by the Neighborhood Tabulation Area[34] (NTA) they belonged to. For each neighborhood, the percentage composition by cluster and its deviation from the Manhattan baseline were calculated (see Appendix for full analysis). The calculated typological deviation of each neighborhood from the Manhattan baseline is coherent with our

[34] https://www1.nyc.gov/site/planning/data-maps/open-data/dwn-nynta.page

local knowledge: Upper West Side resembles the Manhattan style most clostly, while Stuyvesant Town-Peter Cooper Village contains modern red brick apartment towers that are independent of the grid system (**Figure x**).[35] Stuyvesant had a uniquely low percentage of cluster 10 layouts and high percentage of cluster 0 layouts, indicating very sparsely distributed buildings and very uneven grain. This reflects the Town's peculiar architecture, relative to the Manhattan baseline (**Figure x**).

Figure x. Deviation of urban typology distributions from the Manhattan baseline.

Figure x. (Left) Aerial view of Stuyvesant Town-Peter Cooper Village[36]; (Right) Aerial View of Upper West Side[37]

6.2.3 Constructing a standard metric for typological signatures

By reducing the multi-dimensional typological signature into a single-value metric, we achieve the research objective of deriving a standard metric for evaluating 2D city layout geometry. Applying PCA, the top principal component explains 96.8% of the total variance. This standard metric faithfully represents the typological signature of the neighborhood. Using this metric, spatial maps of the top principal components and their deviation from the Manhattan baseline were visualized for each NTA, clearly exposing spatial outliers like Stuyvesant Town and Manhattanville (**Figure x**).

[35] Mumford, E. (1995). 'The "tower in a park" in America: Theory and practice, 1920-1960', *Planning Perspectives*, 10(1), pp. 17-41.
[36] Retrieved from Bloomberg, 2015.
[37] Retrieved from YouTube, 2017.

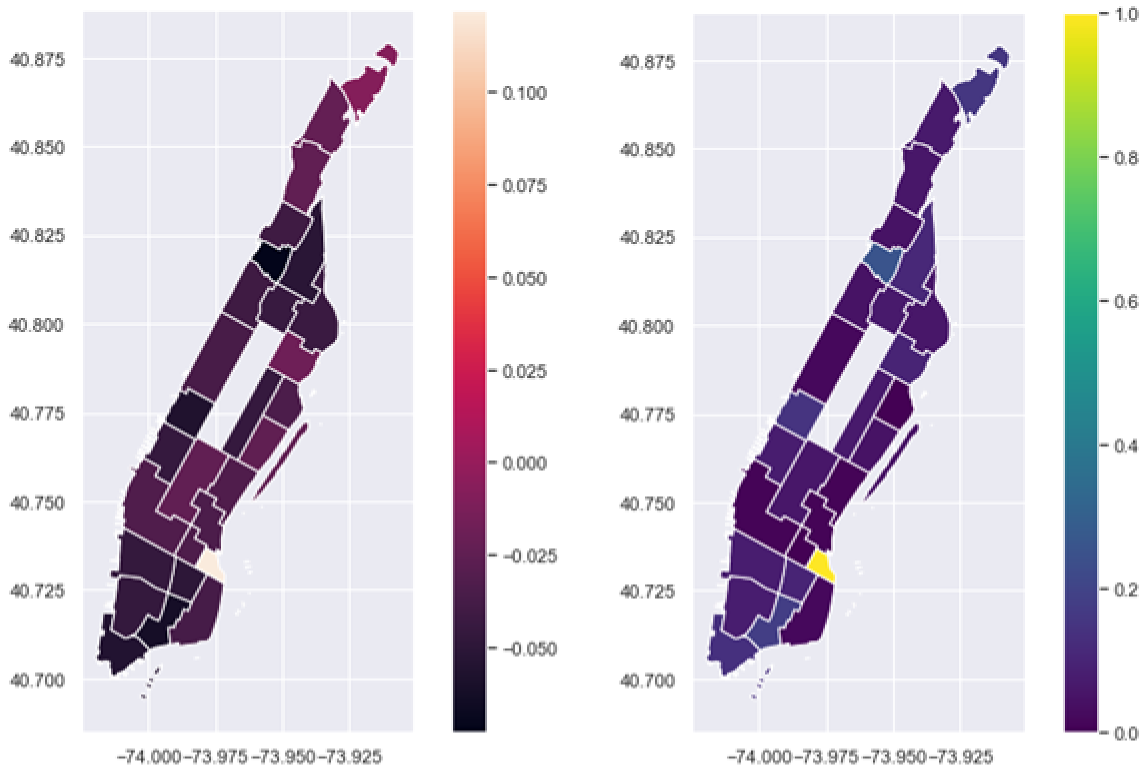


Figure x. (left) Classification of NTAs by top principal component value representing neighborhood typological signature; (right) Classification of NTAs by deviation from baseline.

6.3 Blending city layouts

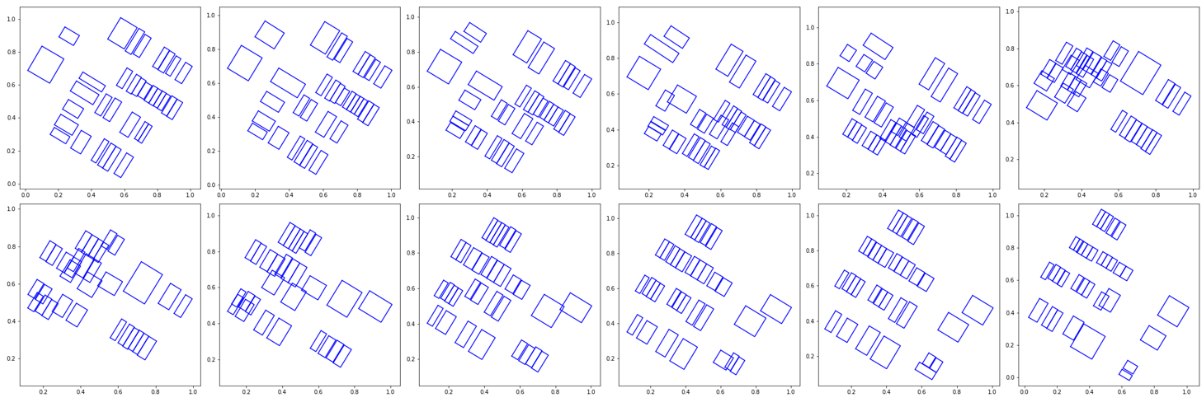
Designing new urban layouts is a tedious and inefficient process. The development area's surrounding urban typology is a key consideration as planners want to control the degree of design consistency relative to the surrounding built environment. Therefore, blending city layouts from surrounding areas could generate contextually sensitive concept layouts.

Although there are commercially available tools that leverage procedural modeling to generate city layouts, they cannot do so without extensive handcrafted procedural rules. Consequently, tools like Delve[38] and City Engine[39] require manual refinement for adaptation to the local context. This is unlike our proposed approach.

Since every city layout is encoded into a shared latent space, the linearly interpolated values of two latent subspaces can be decoded using AETree to generate new city layouts that bear resemblance to both original layouts. **Figure x** demonstrates the blending of city layouts from Clinton and Lower Manhattan, a process which takes merely seconds to generate. Although there are imperfections such as overlapping boxes, we believe that further tuning of the models could generate realistic composite city layouts.

Figure x. Blending of city layouts from Clinton (top-left) and Lower Manhattan (bottom-right).

[38] https://www.sidewalklabs.com/products/delve
[39] https://www.esri.com/en-us/arcgis/products/arcgis-cityengine/overview



6.4 Limitations [YL]

7. Conclusion

The future of urban planning is in computational design, requiring data-driven methods to encode city geometry for new urban planning applications. Rapid urbanization demands scalable ways to generate well-designed city layouts. But this requires a coherent way of describing the complex geometry of city layouts - even before we normatively evaluate whether a particular layout is appropriately designed.

Prior work sought to manually enumerate basic geometric features, failing to capture the diversity and complexity of urban typologies. By leveraging recent advancements in deep learning, we propose a standard metric that faithfully represents the spatial structure of city layouts. Urban planners can use this standard metric to classify, measure and generate new city layouts. We showcase the utility of this standard metric in three proof-of-concept urban planning applications: (1) classifying urban typologies, (2) measuring neighborhood geometries, (3) generating composite layouts.

7.1 Novelty Statement

The contributions of our research are as follows:

- Introducing an LSTM encoding process that learns the spatial information of city layouts and outputs a hidden state vector that is statistically representative of basic urban geometric features.

- Demonstrating quantitatively and qualitatively that the principal components of the encoded latent space remain faithful representations of urban geometry.

- Adapting GMM clustering algorithms to segment the higher-dimensional latent space into distinct urban typologies.

- Introducing an urban typological signature to define the urban geometry profile of neighborhoods and cities.

- Demonstrating the generation of composite city layouts by linearly interpolating two distinct latent subspaces.

- Developing an interactive web application to demonstrate planning applications: https://share.streamlit.io/lazfishing/manhattan-city-layouts

7.2 Future Work

Overcoming the limitations described in Sections 4.4 and 6.4 will involve further model refinement and experiments. Some potential architectural improvements have been explored in the Appendix. Importantly, future work should consider encoding non-orthogonal building footprints and building heights as spatial information, to capture the rich complexity of urban geometry.

Deeper integration into the urban planning workflow through pilots with planning studios and municipal agencies is another important step. This could take the form of a city-wide urban geometry study or a comparative evaluation among cities in the US where building footprint data is widely available.[40]

[40] Heris, M. P., Foks, N. L., Bagstad, K. J., Troy, A. & Ancona, Z. H. (2020). 'A rasterized building footprint dataset for the United States', *Scientific Data*, 7, 207.

APPENDIX

3.4.1 64-Box City Layout Generation

Using a similar LSTMCell architecture, we trained the existing AETree model on larger 64-box city layouts. We adjusted the model hyperparameters including the learning rate and loss calculations. Although our results have improved significantly since the Second Progress Report, there are still obvious spatial discrepancies between ground truth and reconstruction (Figure 5).

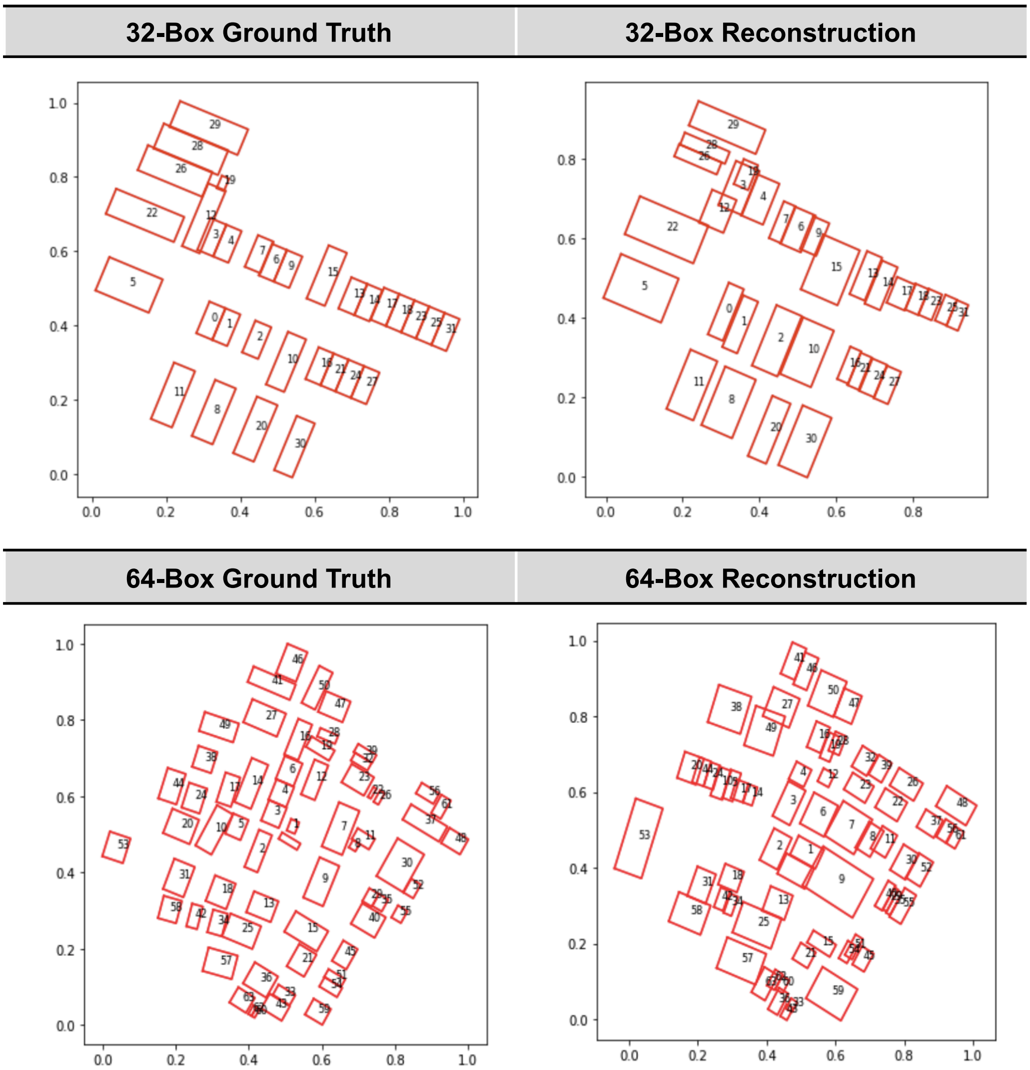


Figure 5. Ground truth and reconstruction of 32/64-box city layouts using LSTMCell.

3.4.2 Transformer Model Substitution

Limitations of the existing LSTMCell recurrent neural network (RNN) include: difficulties in learning long-range dependencies between distant features, computationally intensive architecture and tendency for information to become corrupted (“vanishing gradient problem”). Attention-based transformer architecture has been suggested as a more efficient and effective alternative to RNNs[41]. The proposed Transformer-AETree architecture consists of a tree encoder containing a single self-attention mechanism, and a decoder to reconstruct the original from the encoded features. However, this technique was unable to reconstruct 2D city layouts accurately (Figure 6). Future work may include increasing the number of encoding and decoding layers, and introducing multi-head self-attention mechanisms. These modifications have resulted in superior learning abilities in sequence-to-sequence natural language tasks[42].

32-Box Ground Truth	32-Box Reconstruction

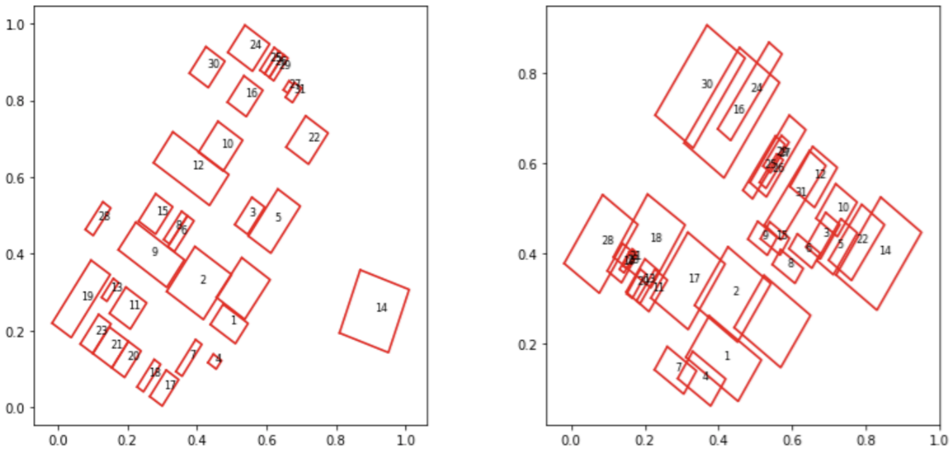


Figure 6. Ground truth and reconstruction of 32-box city layouts using the Transformer model.

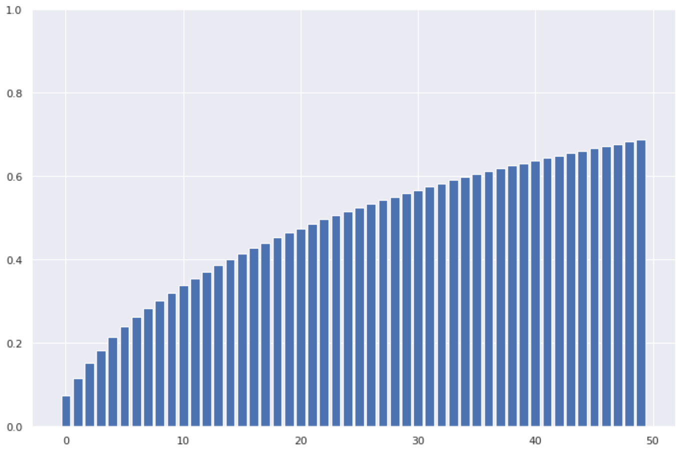


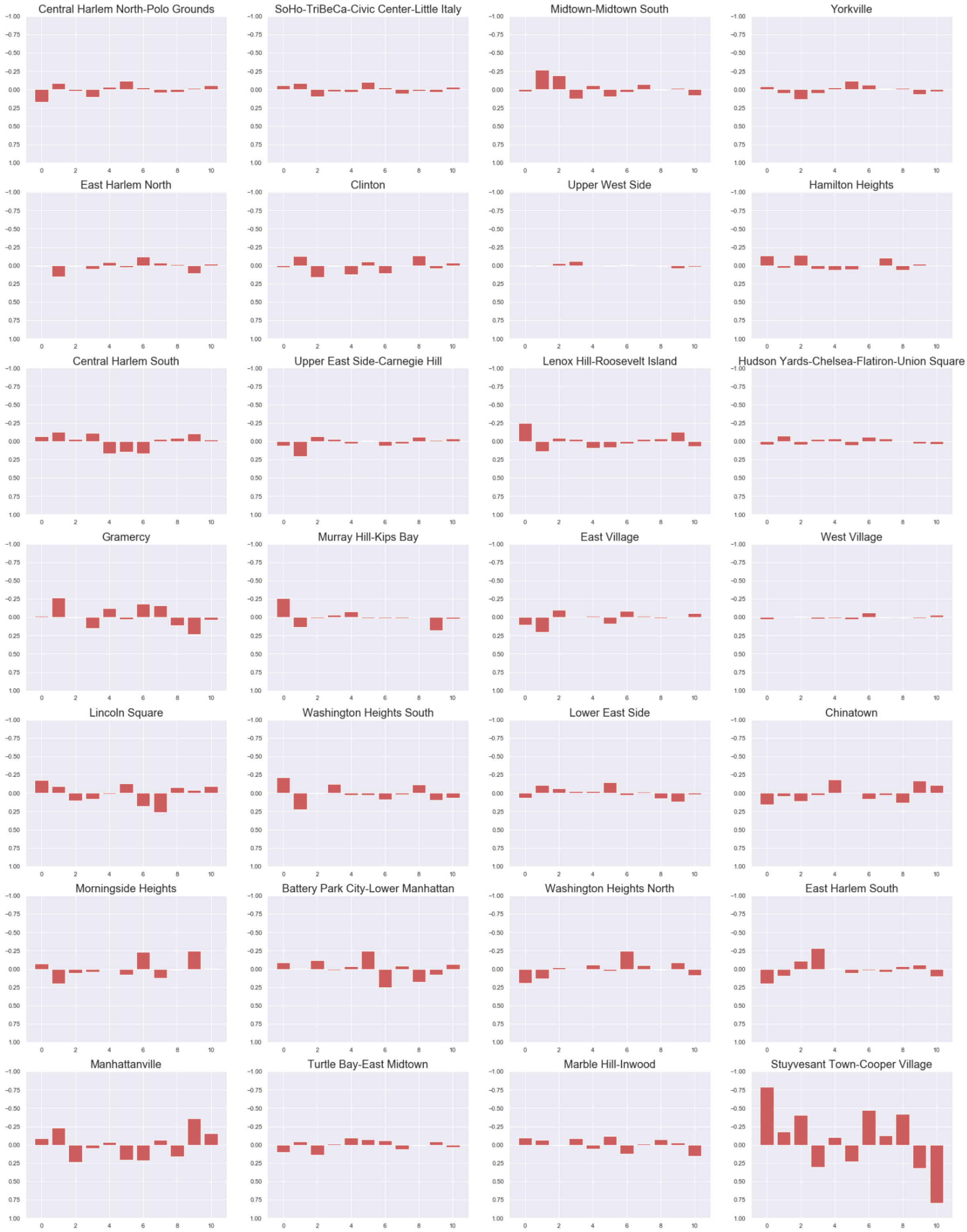
Figure X: Cumulative sum of explained variance.

---

[41] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017). 'Attention is all you need'. In proceedings: 31st Conference on Neural Information Processing Systems (NIPS 2017).

[42] Ibid., 9.

Cluster number	**Density** (% area covered by building footprints)	**Building Regularity** (average elongation of buildings, or length:width ratio)	**Evenness** (size and shape similarity of buildings)	**Orientation** (whether buildings have consistent orientation)
0	Very sparse	Mostly regular footprints	Uneven grain	Mixed orientation
1	Very dense	Irregular footprints	Very even grain	Consistent orientation
2	Sparse	Irregular footprints	Uneven grain	Mostly consistent orientation
3	Dense	Regular footprints	Mostly even grain	Mostly consistent orientation
4	Moderate density	Regular footprints	Mostly even grain	Mostly consistent orientation
5	Moderate density	Regular footprints	Mostly even grain	Mixed orientation
6	Very sparse	Mostly regular foorptints	Very even grain	Mixed orientation
7	Dense	Regular footprints	Uneven grain	Mostly consistent orientation
8	Moderate density	Regular footprints	Mostly even grain	Consistent orientation
9	Sparse	Irregular footprints	Mostly even grain	Mostly consistent orientation
10	Moderate density	Mostly regular footprints	Very even grain	Consistent orientation


Central Harlem North-Polo Grounds
SoHo-TriBeCa-Civic Center-Little Italy
Midtown-Midtown South
Yorkville
East Harlem North
Clinton
Upper West Side
Hamilton Heights
Central Harlem South
Upper East Side-Carnegie Hill
Lenox Hill-Roosevelt Island
Hudson Yards-Chelsea-Flatiron-Union Square
Gramercy
Murray Hill-Kips Bay
East Village
West Village
Lincoln Square
Washington Heights South
Lower East Side
Chinatown
Morningside Heights
Battery Park City-Lower Manhattan
Washington Heights North
East Harlem South
Manhattanville
Turtle Bay-East Midtown
Marble Hill-Inwood
Stuyvesant Town-Cooper Village

Geometric Feature – Average Length and Width					
s/n	ML Method	R2 Score	MSE*	RSME*	MAE*
	Baseline 1 - ZeroR	0.000	0.255	0.505	0.398
	Baseline 2 – Linear Regression	0.129	1.311	1.145	1.122
1.	Decision Tree	0.559	0.117	0.342	0.265
2.	Random Forest	0.698	0.080	0.283	0.215
3.	SVM	0.809	0.051	0.225	0.168
4.	KNN	0.495	0.134	0.366	0.280

* The lower the score, the better the fit of the model

Figure X. Model performance for average length / width ratio prediction using latent space

Geometric Feature – Standard Deviation of Length and Width					
s/n	ML Method	R2 Score	MSE*	RSME*	MAE*
	Baseline 1 - ZeroR	0.000	0.158	0.397	0.312
	Baseline 2 – Linear Regression	0.064	1.011	1.005	0.960
1.	Decision Tree	0.306	0.108	0.329	0.249
2.	Random Forest	0.391	0.088	0.297	0.228
3.	SVM	0.480	0.081	0.285	0.220
4.	KNN	0.296	0.110	0.331	0.256

* The lower the score, the better the fit of the model

Figure X. Model performance for standard deviation of length / width ratio prediction using latent space

Geometric Feature – Range of Length and Width					
s/n	ML Method	R2 Score	MSE*	RSME*	MAE*
	Baseline 1 - ZeroR	0.000	1.906	1.381	0.991
	Baseline 2 – Linear Regression	0.041	13.625	3.691	3.503
1.	Decision Tree	0.224	1.389	1.179	0.891
2.	Random Forest	0.320	1.216	1.103	0.833
3.	SVM	0.343	1.175	1.084	0.832
4.	KNN	0.219	1.398	1.182	0.890

* The lower the score, the better the fit of the model

Figure X. Model performance for range of length / width ratio prediction using latent space

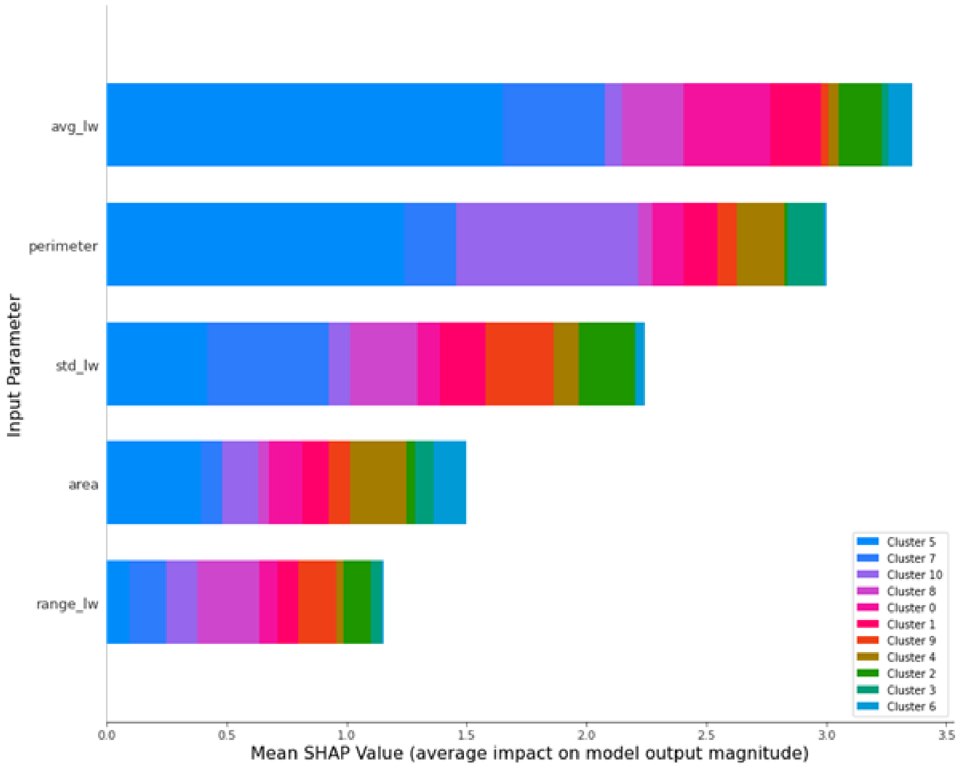


Figure X. SHAP values for the top contributing geometric features in the Logistic Regression model.

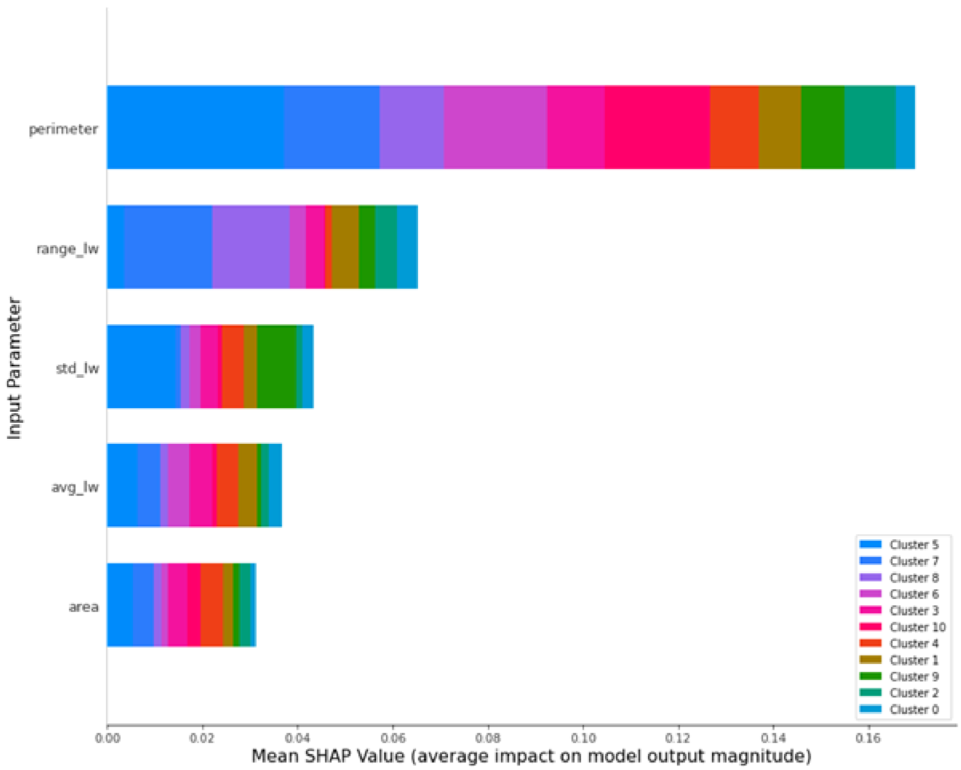


Figure X. SHAP values for the top contributing geometric features in the Decision Tree model.

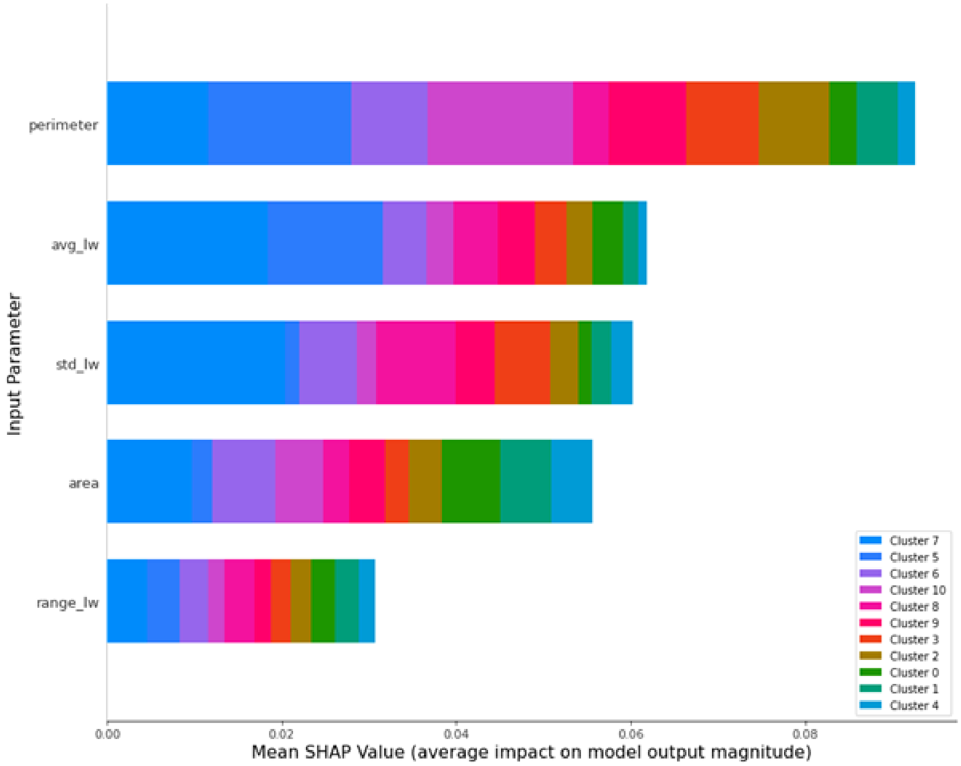


Figure X. SHAP values for the top contributing geometric features in the Random Forest model.

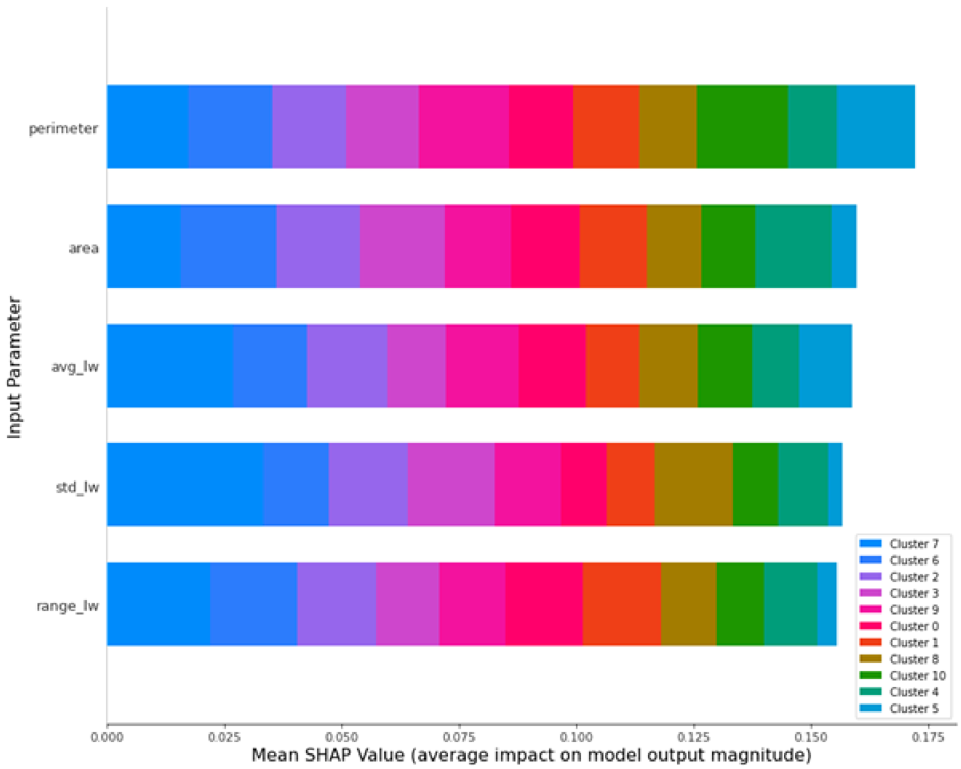


Figure X. SHAP values for the top contributing geometric features in the KNN model.