ISSN: 1648-6897 / eISSN: 1822-4199

VILNIUS TECH
Vilnius Gediminas Technical University

JOURNAL of ENVIRONMENTAL ENGINEERING & LANDSCAPE MANAGEMENT

2026
Volume 34
Issue 2
Pages 171–185
https://doi.org/10.3846/jeelm.2026.26748

INTELLIGENT LAYOUT AND SPATIAL PLANNING OF GREEN URBAN BUILDINGS BASED ON DEEP LEARNING

Yong SUN, Zhengjia XU✉

School of Architecture and Civil Engineering, West Anhui University, 237012 Luan, Anhui, China

Highlights:

- to develop an intelligent layout system using DeepClusGAN to enhance spatial and functional efficiency in urban building arrangements, ensuring building density and urban plans;
- to analyze spatial features and extract key relationships and patterns from high-precision vector data, a Deep Convolutional Generative Adversarial Network is applied and ensures diversity;
- to improve clustering accuracy and adaptability of intelligent layout, the Deep Embedding Cluster model is applied and provides robust building cluster cohesion and separation across urban layout scenarios.

Article History:

- received 27 February 2025
- accepted 27 October 2025

Abstract. The research proposes a novel DeepClusGAN, namely an intelligent combination of Deep Convolutional Generative Adversarial Network (DCGAN) and deep embedding cluster (DEC) to optimize urban building cluster layouts. DCGAN generates innovative and intelligent spatial layouts by learning patterns from existing urban residential building data sources and guarantees diversity and adherence to urban design principles. DEC analyzes the learned spatial features from the previous phase and performs clustering to identify functional zones and spatial relationships between metrics like proximity, compactness, and accessibility. The research maintains a spatial pattern and adjacency of building arrangements by optimizing objectives to meet urban planning goals, such as maximizing space usage, enhancing urban environmental growth, and promoting building architectural harmony. The research achieves an improved silhouette score, an adjusted rand index for validating the building cluster, and a 20.27% improvement in spatial diversity compared to existing models while maintaining an enhanced optimal building layout efficiency.

Keywords: spatial plan, urban building, cluster quality, layout optimization, deep embedding, generative adversarial network.

✉Corresponding author. E-mail: ***zhengjia_xu@hotmail.com***

1. Introduction

As cities grow in size and population, the variety and complexity of urban layouts are more noticeable than ever before, having far-reaching effects on city planning and layout and the quality of life for city dwellers (Shen et al., 2024). Building clustering is essential for large-scale map synthesis, but manual feature and distance metric selection limits its ability to capture urban building complexity and diversity (Sun et al., 2024). Advancements in AI enable an intelligent way of planning the spatial layout of urban buildings that streamlines the existing repetitive and labour-intensive work (Ko et al., 2023). The lack of high-quality residential community layout planning datasets stems from challenges in large-scale real-world data acquisition and time-intensive expert validation (Chen et al., 2023). The building interaction significantly reflects the physical and geometrical characteristics of the urban region, like building floor area, height, and footprints (Ali-Fakulti & Jamil, 2024).

The relationship between spatial patterns and urban clusters can mitigate the regional disparities by analyzing building features based on their structural and functional values (Zhang et al., 2023). The density-based clustering algorithms rely only on homogenous clusters, and their usage in multi-density areas offers inaccurate outcomes in analyzing the building layout (Cesario et al., 2023). Significant research focuses on urban density and building morphological patterns, but the layout through spatial domains is still missing (Formolli et al., 2023). The uneven covering of spatial coordinates biases is analyzed using a machine learning model to investigate the urban agglomerations based on building footprints from Open Street Map (OSM) data (Herfort et al., 2023). The Spatial Planning attributes such as building density, height, floors, and height variance are influence factors (Lan et al., 2022) in predicting urban morphological patterns and neighbourhood layouts.

The research integrates deep learning with regularization techniques to enhance accuracy, reliability, and

interoperability in urban building planning applications, offering probabilistic insights and visual analysis for better decision-making (Wu et al., 2024). The deep learning model utilizes data from geographic information systems, building usage statistics, and environmental conditions to inform optimal spatial layouts (Kalliomäki et al., 2024). Accurate building height data is crucial for assessing building heights' spatial variability and connection (Usui, 2024). Homogeneous-based clustering of inter-buildings to capture spatial relationships and patterns without depending on arbitrary spatial units provides accurate urban density measurement (Montero et al., 2023). Clustering ensemble framework for urban from topologies provides reliable analysis that reveals patterns that enhance urban layout planning (Li & Li, 2024). The Deep Embedding Cluster (DEC) is applied to identify non-linear relationships between urban input attributes like images and categorical inputs, offering cluster stability (de Kok et al., 2024).

Key Contributions are stated in three phases.

1) To develop an intelligent layout system using DeepClusGAN to enhance spatial and functional efficiency in urban building arrangements, ensuring building density and urban plans.
2) To analyze spatial features and extract key relationships and patterns from high-precision vector data, a Deep Convolutional Generative Adversarial Network is applied and ensures diversity.
3) To improve clustering accuracy and adaptability of intelligent layout, the Deep Embedding Cluster model is applied and provides robust building cluster cohesion and separation across urban layout scenarios.

Figure 1 outlines the steps of the research process involved in urban building clusters and spatial planning for intelligent layout.

2. Background studies

Hu et al. (2023) created optimal residential building designs that balance sky view factor (SVF), sunshine duration, and noise. This research combines parametric design (PD), artificial neural networks (ANNs), and multi-objective optimization (MOO). The framework employs ANN models and optimizes using the ideal point technique. Although this method improves design efficiency, it may only apply to specific environments due to its need for detailed data and the difficulty of multi-dimensional optimization.

Vera et al. (2022) examined their responsiveness to geographical resolution, precision, and explanatory features, and this study contrasts Deep Modularity Networks (DMONs) with Gaussian Mixture Models (GMMs) for urban clustering. The significance of social factors in clustering was demonstrated when DMON marginally beat GMM using urban data gathered from a poll of citizens of Santiago, Chile. Possible biases in survey assessment and reliance on input attribute quality for clustering efficacy are limitations.

Technologies like Convolutional Neural Networks (CNNs) automate image analysis within Building Environment research, with data extracted from Google Street View images and social media platforms like Flickr. Yang et al. (2022) used the VGG-16 deep learning architecture, and Wang et al. (2023) learned about the physical characteristics of the spatial planning of urban buildings using the DeepLabv3 model, gathering contextual data from images from street view to estimate the built setting.

Yang et al. (2024) examined geographic data like facility locations and urban demographics. This study utilizes a backpropagation (BP) neural network to maximize commercial building space layouts in Weinan City based on point-of-interest data. The results provide actionable methods for development axes and aggregation centers, demonstrating multi-centred agglomeration patterns and external growth. The difficulty lies in applying the findings to varied urban contexts and its reliance on unique urban datasets.

Zheng et al. (2023) offered a graph-neural network-based reinforcement learning (GNN-RL) model for city planners to use in dealing with the difficulties posed by diverse and irregular topographies. By producing efficient spatial designs customized to different needs, the model surpasses humanized plans in virtual and physical

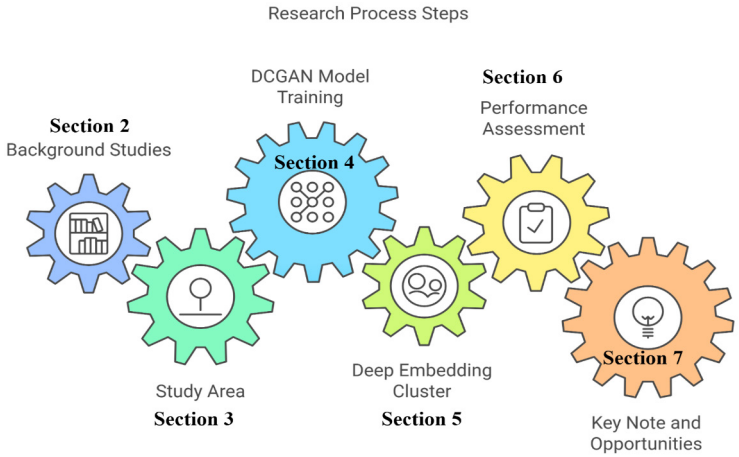


Figure 1. Research process outline

communities and provides better computational performance in less time. The difficulty in applying the model to cities with a wide variety of urban clusters and its dependence on specific datasets are two of its significant drawbacks.

Hao and Wang (2024) presented the graph-based representations of urban systems to forecast the spatiotemporal Interaction (STI) flows between organizational clusters and residential clusters in urban trade districts using a Graph Attention Network (GAT). It was tested using eXplainable AI approaches and developed utilizing data from Miami, Florida. It efficiently captures interdependencies and fluctuating cluster attractiveness. Drawbacks include that it relies on high-quality data and may have trouble scaling to bigger and more complicated urban areas.

Cao and Weng (2024) proposed Sentinel-1/2 pictures and a deep learning-based super-resolution approach. The study estimates building heights with a 2.5 m resolution with RMSEs of 10.318 m in China, 5.654 m in the CONUS, and 4.113 m in Europe. A dataset of 45,000 observations for urban study showcases China's tallest structures. Dependence upon Sentinel data and difficulties with regional generalization are limitations.

Lu et al. (2023) examined the elements that impact the energy consumption of commercial buildings in Singapore, considering building information, behaviors, and the urban environment as factors. The accuracy of predictions is enhanced by 38% with the use of spatially weighted regression, K-means clustering, and artificial neural networks. Results show that land surface temperature has the most significant adverse effect on building energy consumption (BEC), whereas building covering area is an essential positive driver, highlighting the importance of spatial heterogeneity and sustainable urban development.

A key research gap identified from an existing model highlights the potential of deep learning in urban planning, building design, and spatial analysis. Still, it faces difficulties with high-quality datasets, limited generalizability across diverse urban contexts, and inefficiencies in capturing spatial heterogeneity. These limitations are addressed in this current DeepClusGAN research model to generate high-quality, diverse urban environments. Cluster enhancement using DEC enables urban planning with improved adaptability to various spatial and building functionalities.

Table 1. Key characteristics of ReCo dataset

Attribute	Description
Datasource name	ReCo Dataset
Data format	JSON and GeoJSON
Number of communities	37,646
Number of buildings	598,728
Average no. of buildings per community	15.90
Maximum no. of buildings in a community	70,614
Minimum no. of buildings in a community	143
Mean no. of communities	627.43
Data source	Real-world data from 60 cities

3. Study area

The Residential Community (ReCo) Layout Planning data contains information for 37,646 residential community building types and urban characteristics across 60 cities. The dataset is based on high-precision vector coordinates rather than raster images, which preserves spatial attributes and allows for flexible export to various spatial data formats.

Table 1 summarizes the key attributes and their values for the ReCo dataset. It provides valuable insights into the research scope for urban building layouts, its spatial plans for associated communities and buildings, and its coordinates.

Figure 2 provides the minimum, maximum, and average values for the no. of communities, building sizes B_s, and their respective averages are displayed in the bar chart (left). Statistically, these values are shown in the box plot (right), which highlights the range and central patterns of the dataset by showing the variability between categories and the mean. The volume represents the size and morphological variety of urban clusters representing the building and spatial features, incorporating characteristics

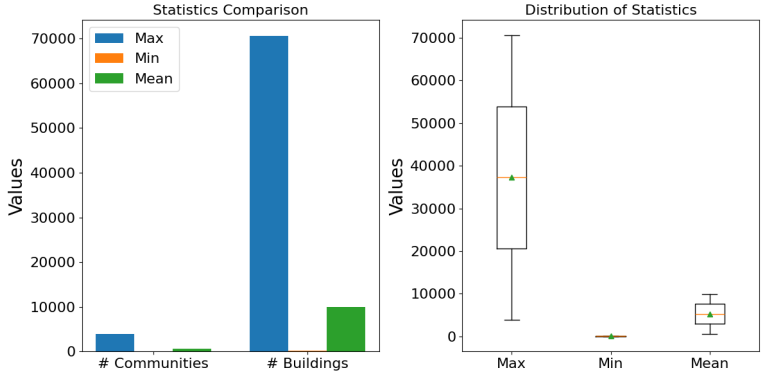


Figure 2. Statistics distribution from the ReCo dataset

like the buildings' solid volume and outside curves.

These diverse properties influence the general features of building types in a given area. The research introduces a novel framework for the intelligent layout of urban building clusters, using deep learning techniques to address spatial efficiency. The spatial features involve proximity ρ_y to reach landmarks like hospitals, schools or transportation. The elevation z influences the weather patterns and infrastructure needs near the considered urban layout. The population density indicates the community CO_D density for determining the urbanization levels, with traffic flow t_f impacts for pollution and accessibility of nearby buildings. The additional green coverage G_{cov} factors near the building include vegetation coverage and air pollution levels. Improving air quality contributes to higher G_{cov} positively to energy efficiency, thermal comfort, and urban liveability.

The $\frac{C_{green} \times \left(1+\ln\left(AQI+1\right)\right)}{1+AQI^2}$ creates a balance where green coverage mitigates the adverse effects of poor air quality.

Figure 3 highlights the building information with ID, floor number, and its spatial structure representation representing the size and shape of the building structure and their complexity in exterior form. The building features represented in Figure 2 involves height H, building area Q_A, volume V, and building age Q_{age} can exhibit combined effects like height-volume interaction or age-area scaling defined using Eq. (1).

$$\varnothing_i\left(H,Q_A,V,Q_{age},d_{landmark},z,CO_D,t_f\right)= \log\left(1+H\times\sqrt{Q_A}\right)+\frac{V^{2/3}}{1+Q_{age}^2}+ \frac{\exp\left(-d_{landmark}/\left(z+1\right)\right)}{1+CO_D}+\frac{\ln\left(1+t_f^2\right)}{\sqrt{d_{landmark}+\int}}. \quad (1)$$

The building $\log\left(1+H\times\sqrt{Q_A}\right)$ represents the logarithmic impact of H and the building area Q_A, whereas $\frac{V^{2/3}}{1+Q_{age}^2}$ normalizes the building volume V by the square of the building age Q_{age}. The spatial features with the consideration of location and accessibility $\frac{\exp\left(-d_{landmark}/\left(z+1\right)\right)}{1+CO_D}$ captures the considering elevation z and community population density CO_D. The parameter $\frac{\ln\left(1+t_f^2\right)}{\sqrt{d_{landmark}+\in}}$ models the logarithmic relationship of traffic flow t_f and its building dependence on $d_{landmark}$. The normalization factors such as $1+CO_D$ and $1+Q_{age}^2$ controls scaling and avoids infinite complexity.

In highly developed areas, where elevation values, building variance, traffic intensity, or contextual factors (COD) can be quite significant, it is important to maintain the composite function's numerical stability. Three design measures make things stable. Normalization and scaling factors are used on each variable $\varnothing_i\left(H,Q_A,V,Q_{age},d_{landmark},z,CO_D,t_f\right)$ before integration. It prevents every feature from taking over the outcome. Second, when logarithmic and exponential operators are utilized, bounded transformations (such $\log(1+x)$ and min–max scaling) are used to smooth the inputs. It is done to avoid problems when working with very big or very small values. Third, the model uses adaptive weighting, which means that different factors, including the age distribution of the population and how close a landmark is to the city, will have a proportional effect on the outcome, even in cities with a wide range of demographics and structures. These mechanisms work together to make sure that the composite function remains stable and gives the same clustering outcomes in different urban environments, such as dense city centers, neighborhoods with a variety of uses, and peri-urban areas that are growing rapidly.

Figure 4 illustrates the overall working procedure of DeepClusGAN for an intelligent building layout plan based on spatial features.

The research idea depicted in Figure 4 demonstrates an intelligent layout for generating urban building layouts, using DCGAN to produce synthetic layouts and then evaluating using the DEC model. The evaluation is analyzed from spatial and building features and then incorporated into the process to represent the cluster cohesion and separation of similar building layouts. Based on their proximity and compactness of urban buildings in the selected communities and contextually generate the appropriate design for urban buildings. In the design of DeepClusGAN (Figure 4), latent vectors $z \sim N(0,1)$ store probabilistic representations of spatial structures derived from training data.

Figure 3. Interactive building layouts in cluster form

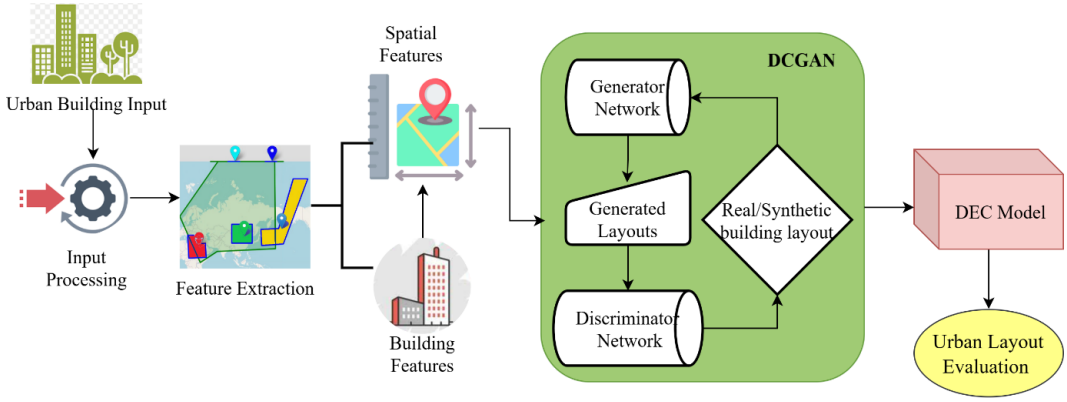


Figure 4. Illustration of DeepClusGAN model for urban building layout

The Gaussian prior gives one a general sample space, but during training, these vectors are molded to hold useful urban patterns like density gradients, orientation alignments, or proximity distributions. This means that different parts of the latent space correlate to different types of urban morphology, which makes it possible to manage the generation. Additionally, by integrating the generator with feature embeddings from DEC, it becomes feasible to correlate latent dimensions with interpretable semantics (e.g., building size, compactness, accessibility), so providing planners with insights into how variations in z translate into observable layout characteristics.

4. DCGAN model training as feature extractor

The training data consists of the spatial and building feature information with community clusters, no. of buildings, and spatial arrangement. The feature collection involves a top-level container with an array of feature objects representing geographic data collection. Each feature object consists of two mandatory properties: geometry, such as point, line-string, or polygon, using coordinates and properties representing an object containing non-spatial attributes associated with the feature, such as name population.

4.1. Generator and discriminator network

The generator L_G learns to capture the underlying spatial patterns, functional zoning, and aesthetic considerations of residential layouts. The generator tries to play a discriminator into classifying its synthetic layout as accurate, where $z \sim p_z(z)$ indicates that the z as random latent vectors sampled from a prior distribution $z \sim \mathcal{N}(0,1)$ into a spatial layout X, represented as polygonal features using Gaussian, and these vectors encode the generator's input noise. The term $G(z)$ generates synthetic residential layouts such as spatial patterns and community building designs using Eq. (2).

$$L_G = -\mathbb{E}_{z\sim p_z(z)}\left[\log D\left(G(z)\right)\right]; \tag{2}$$

$$L_D = \begin{cases} -\mathbb{E}_{x\sim p_{data}(x)}\left[\log D(x)\right] - \mathbb{E}_{z\sim p_z(z)}\left[\log\left(1-D\left(G(z)\right)\right)\right] \\ such\,that \begin{cases} if\ D(x) \approx 1 \text{ real layouts} \\ if\ D(x) \approx 0 \text{ synthetic layouts} \end{cases} \end{cases}. \tag{3}$$

The parameter $D\left(G(z)\right)$ defines the probability assigned by the discriminator L_D that $\left(G(z)\right)$ is real. Here, the generator is penalized when the discriminator successfully identifies $\left(G(z)\right)$ as the fake layout of the building pattern and minimizing the $\log D\left(G(z)\right)$ encouraging it to produce layouts indistinguishable from accurate data. Here $x \sim p_{data}(x)$ as accurate residential layouts from the ReCo urban dataset and $D(x) \approx 0$ in Eq. (3) says the probability assigned to accurate layouts by the discriminator, and $1-D\left(G(z)\right)$ indicates the probability assigned to synthetic layouts as fake.

The amount of spatial variability that the model can capture is determined by the generator latent dimension. Diverse urban morphologies collapse into similar layouts if z is too small, which reduces cluster separation. If it's too large, the model could produce erratic or noisy patterns that divide groups out. Stable and comprehensible clustering is balanced with diversity in moderate levels.

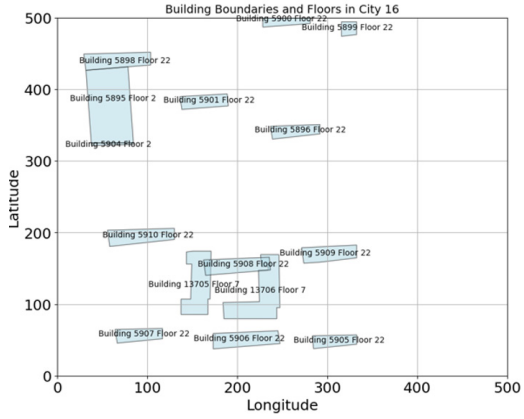


Figure 5. Building boundaries of urban layouts

Figure 5 explores the spatial representation of urban layouts using the polygonal mapping technique and contributes to urban planning by plotting multi-storey building distributions and layouts. This polygonal mapping involves the spatial areas as interconnected polygons, defined by vertices and edges connected using an improved k-means clustering through DEC. The building layouts for sequential 5905, 5906, and 5907 represent the same floor 22 as a cluster form by considering city 16 as a sample. This method analyzes it from GIS data integration for dynamic urban clusters to make better decision-making for effective layouts and modern urban landscapes.

4.2. Adversarial training loop

The generator aims to produce layouts that can fool the discriminator to improve its stability to detect fake layouts, creating a feedback loop that enhances both networks. Let the spatial layout be represented as a feature matrix $X \in \mathbb{R}^{N\times D}$, where N is the no. of polygons, and D represents features like vertex coordinates and area. The neural encoder $f_\theta : X \to Z_{\dim}$ maps X into a latent space $Z_{dim} \in \mathbb{R}^{N\times d}$, where d"D. The embedding Z_{dim} is optimized to preserve geometric and topological relationships between polygons.

Spatial analysis

Density ϑ metrics identify cities with the most densely packed communities to focus on resource allocation and calculate the floor area ratio f_{ar} from building footprint. The spatial interactions include building influence radius BIR and community efficiency index CEI given as given in Eq. (4). The BIR defines the average influence of the radius of buildings based on their proximity ρ_y represents to other buildings. This CEI evaluates the efficiency of land use in terms of building coverage and open space.

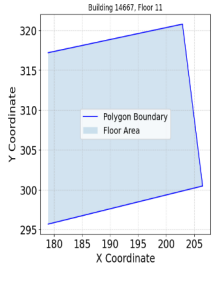
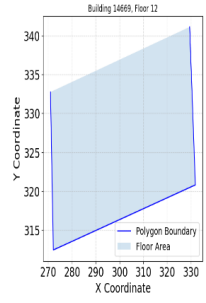
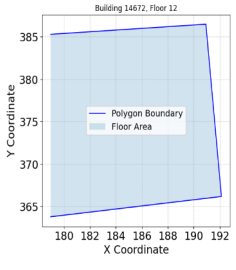
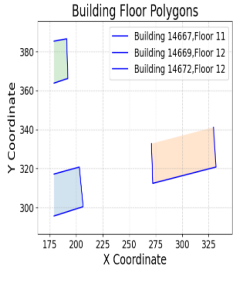
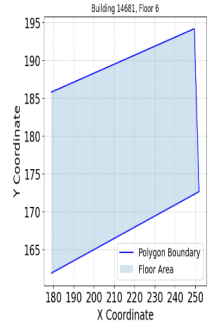
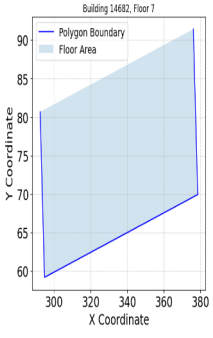
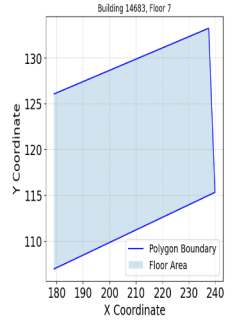
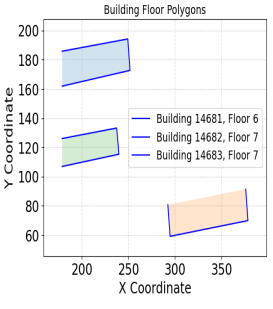
$$BIR = \frac{\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} \exp\left(-\frac{D_{ij}^2}{2\sigma^2}\right)}{N}, \quad CEI = \frac{\sum_{i=1}^{N_b} A_i}{A_c}. \tag{4}$$

Wherein σ defines the scaling parameter that controls the influence decay rate and D_{ij} defines the distance between building i and j. The area of the i^{th} the building footprint is termed as A_i among the total area of the community A_c.

Table 2 defines the spatial coordinates in the form of X and Y of buildings with the represented floor area and polygon boundaries to analyze the floors. The spatial diversity index S_{div} is used to calculate the heterogeneity of building types within a community. The optimized residential layout opt_{rl} efficiency metric can be given using Eq. (5) as follows.

$$opt_{rl} = \alpha \times \frac{\sum_{i=1}^{Q}\left(L_i \times W_i \times H_i\right)}{A_c} + \beta \times S_{CI} + \gamma \times \frac{1}{Q}\sum_{m=1}^{Q}\sum_{n=1}^{R}\frac{1}{d_{mn}} + \delta \times \frac{\sum_{p=1}^{Q}\sum_{q=1,q\neq p}^{Q} w_{ij}}{Q(Q-1)} - \varepsilon \times \underbrace{\left(-\sum_{z=1}^{U}\left(\frac{T_z}{Q}\right)\ln\left(\frac{T_z}{Q}\right)\right)}_{S_{div}} + \zeta \times \frac{\sum_{i=1}^{Q}\sum_{j=1,j\neq i}^{Q}\exp\left(-\frac{D_{ij}^2}{2\sigma^2}\right)}{Q} + \eta \times \frac{\sum_{k=1}^{Q} L_k}{A_c}. \tag{5}$$

Table 2. Geometric representations of building layouts floor and interior plans

Input: Building IDs, Floor Number, and Coordinates			Polygon Boundary
			
			

The total no. of buildings Q in the dataset $\left(L_i \times W_i \times H_i\right)$ represents the length, width, and height of the building i can be given as B_{H_i}, and Φ perimeter of the urban community by calculating the urban density metric using this $\frac{\sum_{i=1}^{Q}\left(L_i \times W_i \times H_i\right)}{A_c}$. The spatial compactness index S_{CI} is given as $\frac{\Phi}{4\sqrt{\pi A_c}}$ with the perimeter Φ of the community boundary (m) with the area of the community A_c measured in terms of m^2. If the value of S_{CI} closer to 1 indicates a highly compact layout. The parameter d_{mn} includes the distance between building m and access point n among total no. of access points R states the amenities like schools, hospitals, and parks from each building R in a community, where T_z defines the count of building type z and the no. of building types U with decay parameter k for the *BIR*. The weight w_{ij} of the edge between building i and j is calculated as $\frac{1}{d_{pq}}$.

DeepClusGAN's discriminator has been optimized to operate effectively in communities with a wide range of building densities, from as little as 143 buildings to more than 70,000. To do this, the discriminator doesn't just look at raw density; it also looks at normalized spatial variables including compactness, proximity distributions, and boundaries geometry. It scale its evaluation across both sparse and dense urban clusters. In low-density environments, where adversarial training can result in mode collapse, two protective measures are implemented: (i) feature normalization guarantees that frequent however legitimate sparse patterns are not erroneously classified as unrealistic, and (ii) diversity-enhancing penalties in the generator loss promote the creation of various layout variants instead of monotonous low-density outputs. Tests on the ReCo dataset show that these approaches help prevent collapse and let the discriminator see realistic layouts at both ends of the urban density range. This makes sure that adversarial training is robust across different community sizes.

When utilizing high-resolution vector polygon data, it is important to keep the geometric precision so that the curves, edges, and vertex alignments remain true to real-world architectural structures. DeepClusGAN solves this problem by adding a geometric regularization term to the adversarial training loop. This limits the generator to making polygons with valid topologies and spatial continuity. Along with the reconstruction loss from the autoencoder, the framework uses structural similarity metrics (SSIM) and polygonal geometric loss functions to check outputs. These functions correct vertices that have not been aligned, borders that are shifted, or polygon closures that are not acceptable. This makes sure that the layouts that are made keep the shape fidelity of each building and the structural coherence of groups. The approach reduces geometric loss while keeping spatial diversity, which means that the synthetic polygons remain true to high-resolution urban datasets without losing realism or computing efficiency.

5. Deep embedding cluster for an intelligent layout

The clustering objective is to group similar spatial zones representing residential and commercial buildings by optimizing a clustering loss function the cluster assignment groups buildings into clusters based on spatial proximity, building type, and functional characteristics. The model optimizes the clustering by minimizing a clustering loss function such as the Kullback-Leibler divergence between a soft assignment distribution and a target distribution.

Table 3. Model configuration and algorithmic parameters

Parameter	Symbol	Value	Description
Learning rate	η	0.001	Step size
Batch size	N	64	No. of polygons processed per iteration
No. of clusters	K	Adaptive	Optimal cluster count
Embedding dimension	d	128	The dimensionality of the latent feature representation
DEC loss weight	α	0.1	Hyperparameter controlling the importance of $\mathcal{L}_{clust}$
BIR parameter	σ	1.0	
Generator Latent dim	Z_{dim}	100	DCGAN input
Discriminator threshold	γ	0.5	Probability cutoff
DEC soft assignment	q_{ik}	Dynamically computed	Latent space proximity
Clustering regularization	λ	0.01	Clustering stability

Table 3 represents the parameters for fine-tuning the integrated DeepClusGAN model with the DCGAN and DEC approach for building layouts based on latent features. The η controls the step size in the optimization process, ensuring efficient convergence. The batch size N indicates how many polygons with the boundary regions are iterated and its impact on training stability and computational efficiency. The adaptive K and d allow for clustering dimensions for balancing the clustering objectives of urban buildings based on their analyzed spatial plans to prioritize the latent space features. Also, the σ influences the clustering spatial organization, while the DCGAN L_G and L_D parameters Z_{dim} and γ controls the generation of synthetic data and enhances urban building planning.

Understanding urban layout features and their impact on building arrangement and clustering provides valuable insights for urban planning and growth. The cluster ranges below indicate that cluster 0 is high-density, small, closely spaced buildings. Cluster 1 is a balanced optimized layout with moderate density and spacing, and Cluster 2 defines the low density, large buildings, and widely spaced urban plan layout.

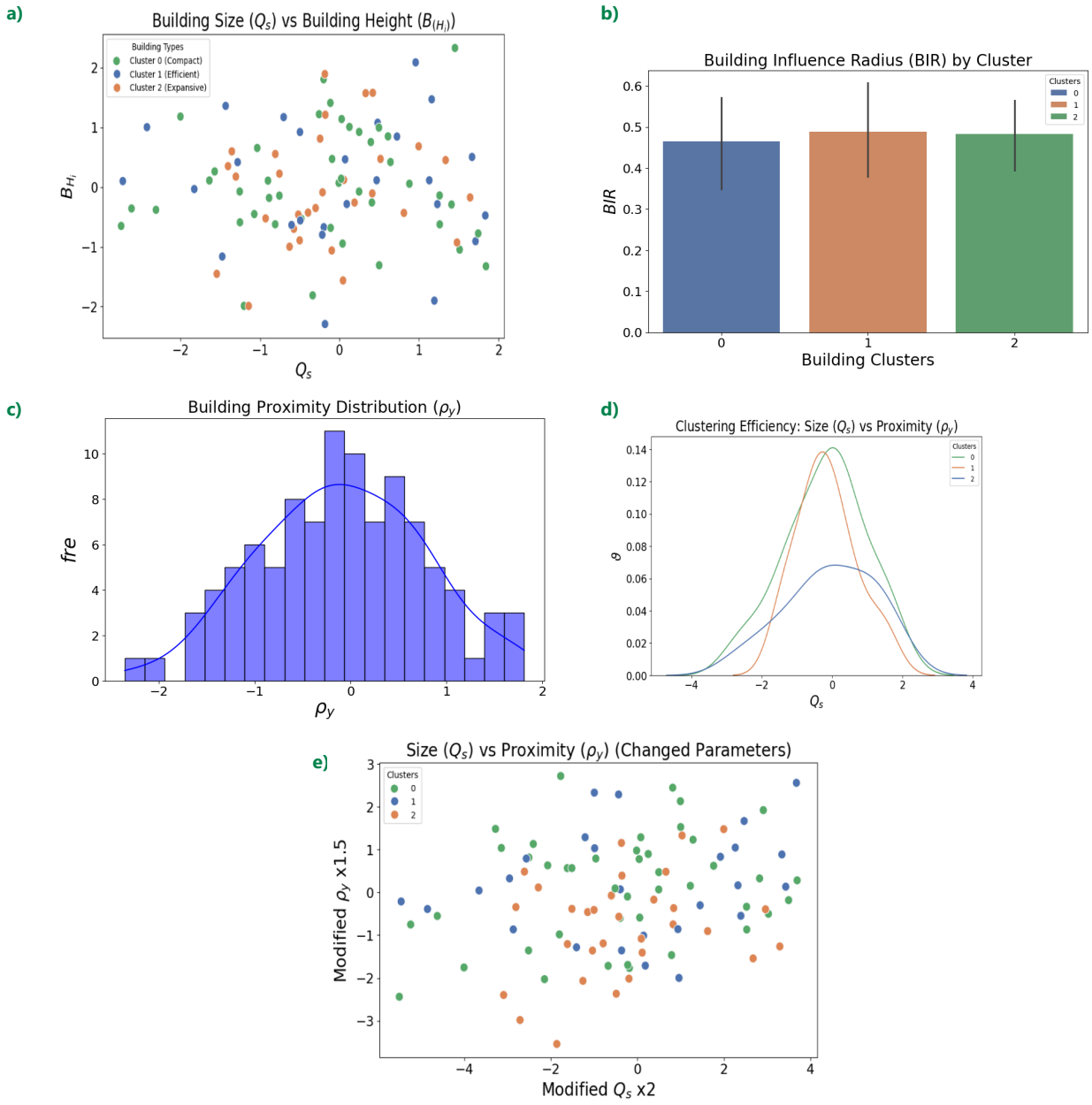


Figure 6. The dimensional feature clustering for urban building layouts: a) Q_s Vs. B_{H_i}; b) *BIR* across clusters; c) Distribution of building ρ_y; d) Clustering efficiency of Q_s vs ρ_y; e) Relationship between Q_s and ρ_y

The plot shown in Figure 6 above demonstrates the relationship between various latent features in building layouts, including size Q_s, density ϑ, proximity range ρ_y and the height of the building is represented as B_{H_i} corresponding to its orientation O, is analyzed across three building clusters: compact, efficient, and expansive. The scatter plot of Q_s Vs. B_{H_i} in Figure 6a examines the correlation between the two attributes within each cluster, whereas Figure 6b compares the *BIR* across clusters, indicating the spatial reachability of buildings. The histogram shown in Figure 6c illustrates the distribution of building ρ_y across the spatial coordinates highlights the cluster tendencies. The ρ_y data visualized using histogram bars and a kernel density estimation curve provides the frequency distribution of the building proximity range. It shows that $\uparrow \rho_y$ high values suggest a more spaced building arrangement, while a lower $\downarrow \rho_y$ represents the buildings closer to each other and clustered under one community. Figure 6d demonstrates the density plot where; it explores the joint distribution of Q_s and ρ_y, providing knowledge for urban planners regarding how these features interact within the clusters. Additionally, Figure 6e represents a scatter plot for modified parameters with a scaled range of varying Q_s and ρ_y indicates how these changes influence the optimal cluster pattern of urban buildings.

When clusters have different densities or are not well separated, using KL divergence as the clustering loss in DEC can lead to issues. Clusters that overlap or are inconsistently sized can be driven into deformed shapes by KL's

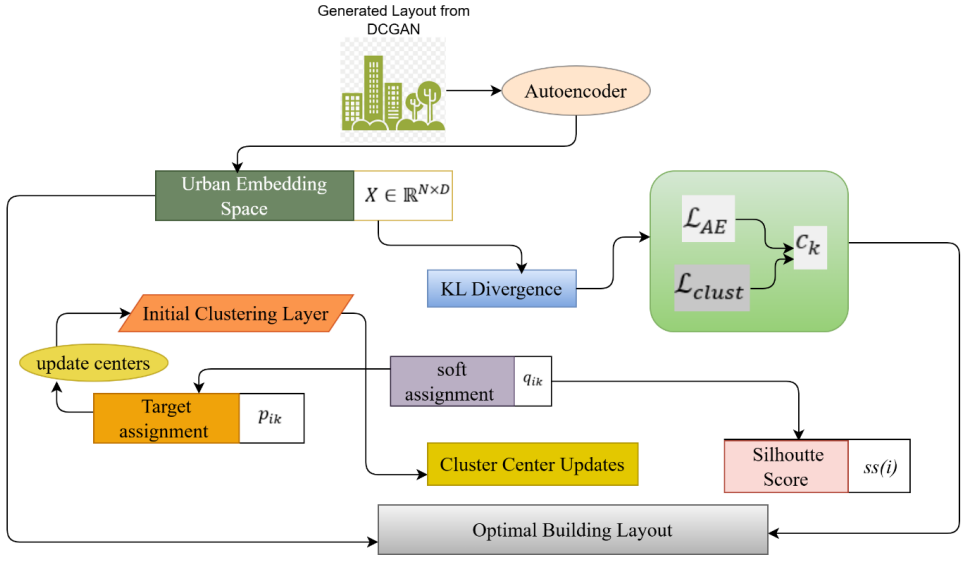


Figure 7. Portray of deep embedding cluster for building layout generation

tendency to sharpen allocations toward cluster centers. This can lower the quality of spatial grouping in dense or irregular urban data by causing incorrect embeddings, unstable cluster boundaries, and misclassification of edge cases.

Figure 7 illustrates that the DCGAN generates the initial spatial structure for layouts. It then encodes the generated layouts into a lower-dimensional urban embedding space and reconstructs them to capture essential building and spatial layout features for urban planning. The initial clustering layer groups similar layouts by analyzing the boundaries with polygon patterns on each floor and their encoded representations. The clustering loss $\mathcal{L}_{clust}$ encourages the clustering algorithm to group similar layouts in a combined form. The KL divergence measures the difference between the soft q_{ik} and target p_{ik} term assignments to refine the cluster centers based on the assigned layouts. The final quality of the clustering is validated using silhouette score $ss(i)$ for attaining optimal building layouts.

DEC loss functions

Let x_i be the input feature vector representing each spatial polygon like vertices, areas, and adjacency from the DCGAN generator, and C be the no. of clusters given for the total X as the feature matrix of polygons. DEC maps these high-dimensional spatial features into a lower-dimensional embedding space while preserving geometric and topological relationships represented in Eq. (6).

$$\mathcal{L}_{AE} = \begin{cases} \sum_{i=1}^{N} x_i - \hat{x}_i^2, \text{where } x_i \in \mathbb{R}^d,\ \hat{x}_i = f_{decoder}\left(f_{encoder}\left(x_i\right)\right) \\ \text{where } L = \mathcal{L}_{AE} + \alpha \mathcal{L}_{clust} \text{ and } \mathcal{L}_{clust} = \sum_{i=1}^{N} q_i - p_i^2 \\ \text{such that } c_k = \dfrac{\sum_{i=1}^{N} q_{ik} x_i}{\sum_{i=1}^{N} q_{ik}} \end{cases} \quad (6)$$

where $\hat{x}_i$ is the reconstructed input, the autoencoder minimizes the reconstruction error, and N indicates the total no. of feature vectors taken from the ReCo dataset as communities. Clustering is applied in the latent space, optimizing the clustering objective with the parameter $f_{decoder}$ and $f_{encoder}$ represents the encoder and decoder functions in the autoencoder that minimize the reconstruction error. The encoder $f_{encoder}$ learns the embedding $z_i = f_\theta\left(x_i\right)$. Across all input vectors x_i ensuring that the embedding space preserves the meaningful features of the given input urban building dataset. In the clustering loss $\mathcal{L}_{clust}$, deep embedding cluster minimizes the clustering loss in the learned embedding space defined as referred to in Eq. (6). The N is the no. of samples, q_i are the soft cluster assignments q_{ik} with probabilities produced by the model and p_{ik} is the target cluster assignment obtained using k-means with q_i is updated iteratively to improve clustering.

The objective function in DEC can be written as a combination of the reconstruction loss $\text{Re}_{\mathcal{L}}$ and the $\mathcal{L}_{clust}$ given as $\sum_{i=1}^{N} q_i - p_i^2$ where α is a hyperparameter controlling the importance of the $\mathcal{L}_{clust}$ relative to the autoencoder $\mathcal{L}_{AE}$ of $\text{Re}_{\mathcal{L}}$. The cluster centers are updated iteratively in the embedding space and can be given as $\dfrac{\sum_{i=1}^{N} q_{ik} x_i}{\sum_{i=1}^{N} q_{ik}}$ for c_k, where it represents the center of cluster k in the embedding space and q_{ik} as the soft assignment probability of sample i to cluster k.

The illustration shown in Figure 8 of applying DEC with a neural network learns a low dimensional data embedding, enabling effective clustering in the embedded space. The latent building layout features represent the category of overall building size in a layout Q_s with density ϑ,

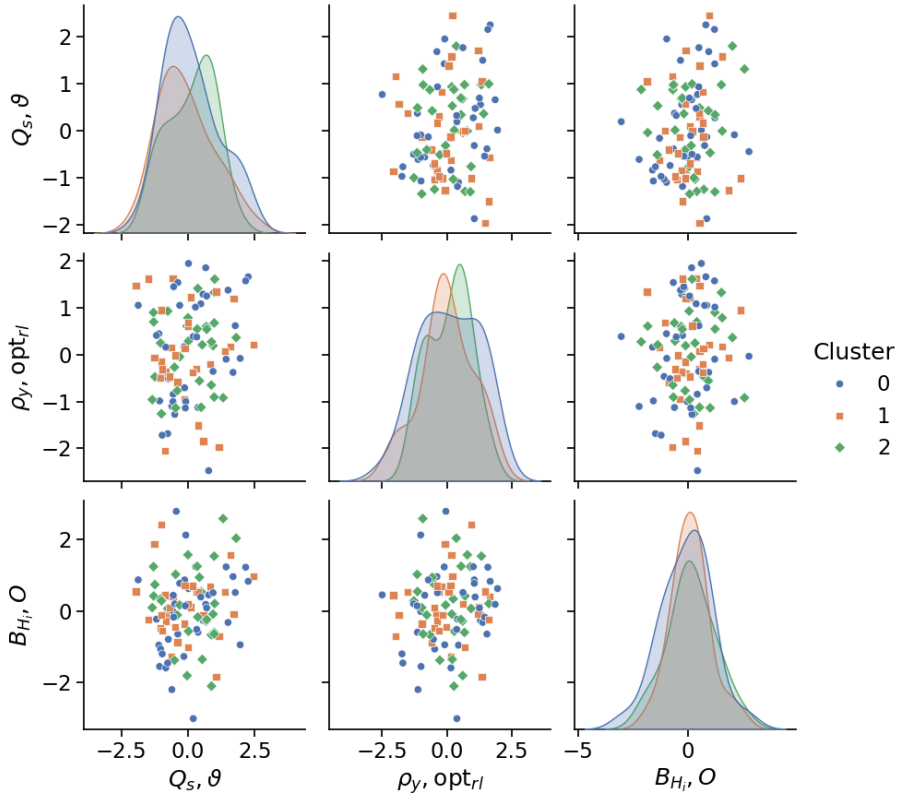


Figure 8. Latent building layout features with cluster types

proximity range ρ_y and optimal residential building layout efficiency can be given as opt_{rl} and the height of the building represented as B_{H_i} corresponding to its orientation. The DEC enables effective clustering in the embedded space and forms clusters corresponding to optimal layout categories termed 0, 1 and 2 for compact, efficient, and expansive based on the spatial coordinates of the building representation.

The soft cluster assignment and target cluster assignment are two types of cluster assignment, including the DEC soft cluster assignment q_{ik} for the i^{th} embedded vector z_i to cluster k is defined with the probability that data point i belongs to cluster k. This soft assignment is dynamically computed based on the distance in the latent space is given as follows using Eq. (7):

$$q_{ik} = \frac{\left(1+z_i - c_k{}^2\right)^{-1}}{\sum_{j=1}^{C}\left(1+z_i - c_k{}^2\right)^{-1}}, \quad \forall k \in \left[1, C\right], \quad \text{where } z_i \in \mathbb{R}^k;\ c_k \in \mathbb{R}^k; \tag{7}$$

$$p_{ik} = \frac{q_{ik}{}^2 / \sum_{i=1}^{N} q_{ik}}{\sum_{j=1}^{C}\left(q_{ij}{}^2 / \sum_{i=1}^{N} q_{ij}\right)}, \tag{8}$$

where $z_i \in \mathbb{R}^k$ defines the feature embedding of sample i, and $c_k \in \mathbb{R}^k$ represents the center of cluster k with C as no. of clusters. The target cluster assignment p_{ik} refines the q_{ik} iteratively using an Eq. (8), with N represents the no. of data points in the dataset and p_{ik} indicates the target probability of data point i belonging to cluster k.

A combination of internal validity indices and spatial performance metrics, including cluster compactness and diversity, is used to find the best number of clusters in an area. These criteria make sure that the clusters are both statistically well-separated and make sense in terms of urban design. In DeepClusGAN, the number of clusters, K, cannot be set in stone. Instead, it evolves as training continues on. It starts with excessive clusters and then merges or divides them based on changes in assignment stability, reconstruction error, and density variation in the latent embedding. Consequently, K converges to the optimal number of clusters that equilibrates cohesiveness and separation among diverse urban morphologies, enabling the model to adapt automatically to communities characterized by significant density variability and structural complexity.

Since urban building data is both statistical and spatial, the model also takes into account problems like spatial autocorrelation and heteroscedasticity that come up when looking at city features that are shown as polygons. Spatial autocorrelation happens when buildings that are close to each other have comparable features, including height, density, or orientation. This makes them observations that are not independent. To solve this problem, the DCGAN generator uses adjacency and neighborhood continuity to make fake layouts. The DEC clustering, on the other hand, puts polygonal characteristics into a latent space that captures both geometric and topological associations. This makes sure that clusters look like real neighborhood blocks instead

than polygons that have been dispersed around. Heteroscedasticity, or the unequal variance of features across different areas, is especially clear when you compare central business districts with a lot of fluctuation to outlying areas that are more consistent. The methodology determines this problem by using normalization and adaptive clustering weights to balance the contributions from areas with high and low variance. The GAN discriminator also punishes implausible variance patterns, which lowers the chance of placing too much weight on extreme values or making building distributions that aren't realistic. DeepClusGAN uses spatial autocorrelation to improve local coherence and eliminate the distortions induced by heteroscedasticity by combining both techniques. The result is architectural layouts that are both statistically valid and natural in context.

The Building Influence Radius (BIR) parameter π defines the geographic reachability scale, which is how much one building affects another during the clustering process. Small π values only focus on nearby neighbors, which might break up dense metropolitan regions into unstable clusters. Large π values, on the other hand, oversmooth the data, combining different neighborhoods and making local details less clear. To avoid these extremes, π is adjusted by looking at typical spatial scales (such the mean distance to the nearest neighbor) or by checking against clustering quality measures like silhouette and ARI. In practice, adaptive or multi-scale settings for π π make cities with mixed densities more durable, making sure that clustering proposals are both locally coherent and regionally significant.

The scaling parameter π^{σ} in the BIR Gaussian kernel determines the rate at which spatial effect decreases with distance. Smaller π^{σ} stresses local cohesiveness, raising compactness and CEI but risking reduced diversity, whereas larger values extend reachability, improving heterogeneity but potentially lowering. It is modified using adaptive learning or data-driven initialization (like nearest-neighbor distances) to find the right balance between these trade-offs.

The DEC framework uses a Student's t-distribution kernel for the soft assignment instead of a Gaussian because the t-distribution has larger tails, which is useful in high-dimensional latent spaces. In reality, this trait makes it less likely that points near the edges of clusters can be forced into unclear memberships, while at the same time making points near the centers of clusters more important. The t-distribution sharpens assignments and increases inter-cluster separation, which makes clusters in urban spatial embeddings more compact and stable. It is better than a Gaussian kernel, which rapidly decreases with distance and can make it challenging to separate clusters in complicated manifolds. The result is that simpler to understand because it makes certain that building designs with varied structural patterns remain separate in the latent space.

6. Performance assessment

The performance assessment of this research is evaluated using TensorFlow GPU, with custom loss functions and hyperparameter tuning for DeepClusGAN optimal performance. For comparative analysis, the existing models, such as DMON-GMM (Vera et al., 2022), GNN-RL (Zheng et al., 2023), and GAT-XAI (Hao & Wang, 2024), are compared with the DeepClusGAN proposed model using evaluation metrics such as cluster quality in terms of silhouette, Adjusted Rand Index (ARI), spatial diversity, and Building Layout Efficiency (BLE) measures.

Transferability to other urban contexts can be diminished if it is overfit to a particular dataset, since it can gather feature noise or variance unique to the dataset instead of actual adaptability. Due to its dependence on inter-model variance across spatial features, spurious diversity can be overestimated. Regularization, cross-dataset validation, and feature selection that prioritizes important spatial qualities (such as density, proximity, and compactness) are used to ensure resilience and ensure that the variation score represents real-world layout flexibility rather than random fluctuations.

Cluster Quality Evaluation

Silhouette score:

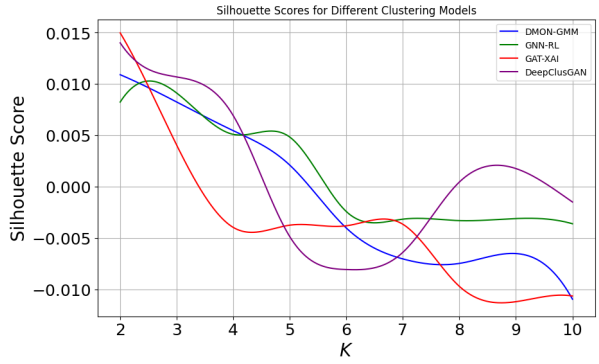


Figure 9. Cluster quality evaluation using $ss(i)$

Figure 9 represents the metric to evaluate the quality of the clustering, which is handled by the silhouette score $ss(i)$ with both the cohesion $a(i)$ called within similarity representing average distance from point i to all other points in the data sample is given as N in the same cluster group. Likewise, the separation $b(i)$ means between cluster dissimilarity in the nearest neighbouring urban building cluster C_j. The $d_{spatial}(i,j)$ represents the spatial distance between buildings i and j, as the $d_{building}(i,j)$ indicates the feature distance with height and energy use as defined in Eq. (9).

$$ss(i)=\begin{cases}\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\dfrac{b(i)-a(i)}{\max\left(a(i),b(i)\right)}, & \text{for each data point } i\\ \text{where, } a(i)=\dfrac{1}{|C_i|-1}\displaystyle\sum_{j\in C_i,\,j\neq i} d_{spatial}(i,j) & .\\ b(i)=\underbrace{\min}_{C_j\neq C_i}\dfrac{1}{|C_j|}\displaystyle\sum_{j\in C_i} d_{building}(i,j)\end{cases} \tag{9}$$

Findings:

The interpretation is higher values of $ss(i) > 1$ indicates well-separated clusters with high cohesion and the lower values of $ss(i) = -1$ indicates poor clustering of buildings with points being incorrectly grouped.

Adjusted Rand Index (ARI)

The Adjusted Rand Index (ARI) quantifies the similarity between two clustering solutions, adjusting the random chance, and it is explained using mathematical representation using Eq. (10) as follows.

$$ARI\,score = \begin{cases} \dfrac{RI - E[RI]}{\max(RI) - E[RI]}, \text{ where} \\ RI = \dfrac{tp + tn}{tp + tn + fp + fn} \text{ and} \\ E[RI] = \dfrac{\sum_{i=1}^{N} \binom{n_i}{2} \binom{N - n_i}{2}}{\binom{N}{2}} \end{cases}. \quad (10)$$

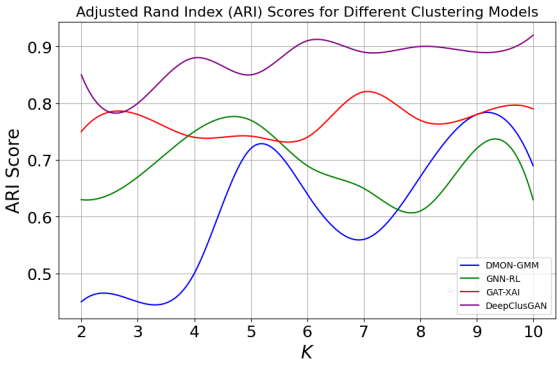


Figure 10. Cluster quality evaluation using *ARI* score

Figure 10 demonstrates that the RI measures the similarity between the predicted clustering and the ground truth, tp defines the no. of pairs of spatial points that are correctly grouped in both the predicted and ground truth clusters. Then, tn defines the no. of spatial pairs of points correctly grouped apart in both clustering, fp represents the no. of pairs of points incorrectly grouped in the predicted clustering. The fn defines the no. of points incorrectly grouped apart in the predicted clustering. The term $E[RI]$ indicates the expected value of the RI for random clustering, and the predicted value accounts for the random chance in building cluster assignments with n_i is the size of the i^{th} cluster in the ground truth, and $\binom{N}{2}$ represents the total no. of unique pairs of points in the embedded building space. The maximum possible value $\max(RI)$ of the RI, which indicates 1, represents the perfect agreement between the predicted and ground truth clustering.

Findings:

Then the $ARI\,score = 1$ indicate that perfect agreement between predicted and ground truth clusters, and $ARI\,score = 0$ suggests random clustering, with no better performance than chance, and the $ARI\,score < 0$ indicates that predicted clustering is worse than random clustering.

Spatial Diversity Index

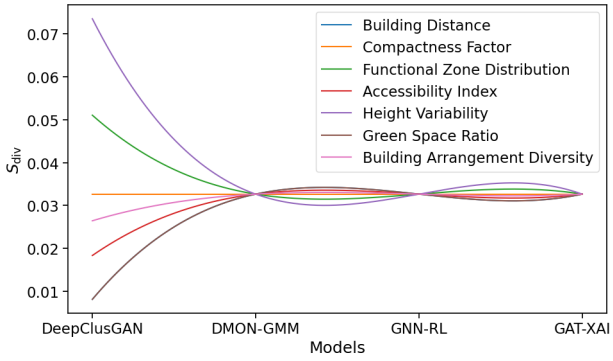


Figure 11. Spatial diversity analysis

Figure 11 demonstrates the spatial diversity index S_{div} depicted in Figure 11 shows that the DeepClusGAN model generates a high degree of flexibility, variety, and creative layouts that maximize S_{div} and promotes variety in spatial-focused building features. The S_{div} is a function of layout features for different models $S_{div} = f(X_1, X_2, \ldots, X_n)$, where X_j is the j^{th} spatial feature for the i^{th} model across n features. This analysis considers the variance of layout features across different models and the derivative regarding feature changes.

$$S_{div} = \begin{cases} \dfrac{1}{n}\sum_{j=1}^{n}\left[\dfrac{1}{M}\sum_{i=1}^{M}\left(X_{ij} - \bar{X}_j\right)^2\right] \\ \text{such that } var(X_j) = \dfrac{1}{M}\sum_{i=1}^{M}\left(X_{ij} - \bar{X}_j\right)^2. \\ \dfrac{\partial S_{div}}{\partial X_j} = \dfrac{2}{nM}\sum_{i=1}^{M}\left(X_{ij} - \bar{X}_j\right) \end{cases} \quad (11)$$

The parameter $\left(X_{ij} - \bar{X}_j\right)$ in Eq. (11) is the difference between the feature value for the i^{th} model and the mean of that feature across all models M. The proposed DeepClusGAN appears to produce layouts with the highest spatial diversity and improves adaptability with priority in urban planning. The DeepClusGAN shows an overall improvement of ~10.27% in S_{div} compared to existing models across all x-axis variants and consistently outperforms the other DMON-GMM (Vera et al., 2022), GNN-RL (Zheng et al., 2023), and GAT-XAI (Hao & Wang, 2024) models.

Building Layout Efficiency

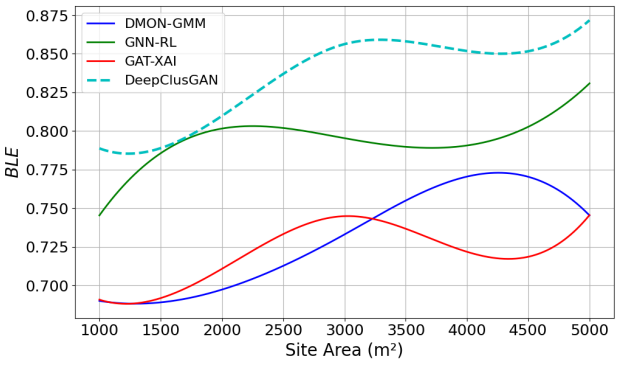


Figure 12. *BLE* comparison analysis

Figure 12 depicts the building site area that represents the total area under consideration, used as an independent variable. This *BLE* metric measures the efficiency in optimizing the layout boundaries for a given area to improve the robustness of the analysis. The analysis measures the effectiveness of the spatial model that optimizes the utilization of boundaries within a given site area and evaluates the clustering efficiency, planning decisions and maximal alignment with urban planning goals.

$$BLE_{opt} = \begin{cases} \underbrace{\arg\min}_{x}\left(\sum_{i=1}^{n}\left(\beta_i \times A_i \times \frac{E_{actual,i}(x)}{E_{optimal,i}(x)}\right)\right) & \text{for each spatial point } i \\ \beta_i = \begin{cases} \alpha_{high}, & \text{if spatial pattern satisfies } H, Q_A, V, Q_{age} \\ \alpha_{low}, & \text{if spatial feature is unsatisfied} \end{cases} \end{cases}, \tag{12}$$

where x in Eq. (12) represents the design variables with layout, green space energy, and environmental surroundings optimized by the DCGAN model. The parameters α_{high} and α_{low} are constants representing the importance of space i in terms of spatial conditions. This layout planning intelligently optimizes the building by minimizing the energy inefficiency between the $\frac{E_{actual,i}(x)}{E_{optimal,i}(x)}$ outcomes represent its importance and significantly contribute to urban planning. By switching on or off contributions from attributes like building height, quality, or vegetation, the binary weighting factor in the BLE formulation provides spatial limitations. Due of the penalties for design that breach environmental or zoning regulations, this changes the optimization landscape. The optimization process can lead to unequal solutions if these criteria are not consistently fulfilled across clusters; some clusters may achieve high efficiency while others decline, which would lower the resultant layouts' overall validity and coherence.

A sensitivity analysis of the clustering weight parameter α in the DEC loss shows how it affects the quality of the embedding and the accuracy of the grouping, especially when the spatial patterns are not regular or are noisy. The parameter α determines how important clustering coherence is compared to reconstruction fidelity. When α is set too low, embeddings tend to focus on reconstruction, which makes polygons look right but makes it hard to separate clusters, especially in urban areas with a lot of different types of people. On the other hand, when α is too high, the model places an excessive emphasis on cluster coherence. It could drive noisy or irregular patterns into fake groups, making it harder to generalize across different city morphologies. Tests on the ReCo dataset reveal that intermediate values of α (0.05–0.2) give the optimal balance, producing stable embeddings that keep both geometric fidelity and cluster consistency. This analysis shows that adjusting α is important for finding the right balance between the capacity to handle noise and grouping urban layouts correctly.

It lets the model balance several goals: maximizing entropy encourages functional diversity and flexibility, while minimizing sprawl and supporting infrastructure efficiency. In real life, high entropy can make congested central business districts less compact, whereas excessive compactness can make suburban layouts monotonous. DeepClusGAN preserves numerical stability and stops over-optimization toward a single goal by altering the weights and using adversarial training. Case studies utilizing the ReCo dataset validate that functions as an adjustable compromise operator, producing layouts that maintain robustness across various urban morphologies while enhancing both adaptability driven by diversity and efficiency driven by compactness.

The relationship between building size and distance between groups can be beneficial for zoning and regulatory decisions as it shows how physical size and distance can affect the shape of a neighborhood. Strong positive correlations may show places where a lot of buildings are too adjacent, which can result in congestion, less livability, and not following zoning rules for spacing. On the other hand, weak or inconsistent correlations may point to parcels not being used to their full potential since they aren't being developed to their zoning potential. Planners can use DEC-based clustering to find areas where present layouts differ from the planned zoning outcomes, such too many business centers in one area or too many residential zones spread out. It can change laws or incentives to fix the problem. So, this combination of correlation analysis and smart clustering gives us a useful tool for finding problems and making urban rules that are more flexible and based on facts.

7. Key note and research opportunities

The introduced DeepClusGAN combines DCGAN and DEC to optimize urban layout planning using the spatial data from 37,646 communities across 60 cities, utilizing vector coordinates to assess building features like density, proximity, and environmental influences. This intelligent model

integrates deep clustering to combine spatial features and applies GAN to generate diverse urban spatial layouts. Here, the spatial features are analyzed using the DCGAN model and clustered with optimal urban layouts for dynamic adaptation. The results demonstrated the model's ability to create balanced urban layouts that enhance the optimized model. The DeepClusGAN model ensures building layout efficiency by optimizing the building arrangements based on spatial relationships with a 10.27% improvement in diversity, with building density and proximity factors. Though the use of the DeepClusGAN system offers benefits, the research has a downside: the selected data source does not fully capture the historical or cultural architectural consideration for layout optimization and makes scalability analysis difficult among large urban areas. Future work will focus on integrating dynamic urban development data for real-time layout optimization. The future development of autonomous planning tools to assist urban planners in designing cities with minimal manual intervention.

Data availability statement

All data generated or analysed during this study are included in this article.

References

Ali-Fakulti, M. F., & Jamil, J. A. (2024). Exploring pattern mining with FCM algorithm for predicting female athlete behaviour in sports analytics. *PatternIQ Mining*, *1*(1), 45–56. https://doi.org/10.70023/piqm245

Cao, Y., & Weng, Q. (2024). A deep learning-based super-resolution method for building height estimation at 2.5 m spatial resolution in the Northern Hemisphere. *Remote Sensing of Environment*, *310*, Article 114241. https://doi.org/10.1016/j.rse.2024.114241

Cesario, E., Lindia, P., & Vinci, A. (2023). Detecting multi-density urban hotspots in a smart city: Approaches, challenges and applications. *Big Data and Cognitive Computing*, *7*(1), Article 29. https://doi.org/10.3390/bdcc7010029

Chen, X., Xiong, Y., Wang, S., Wang, H., Sheng, T., Zhang, Y., & Ye, Y. (2023, October). ReCo: A dataset for residential community layout planning. In *Proceedings of the 31st ACM International Conference on Multimedia* (pp. 397–405). ACM. https://doi.org/10.1145/3581783.3612465

de Kok, J. W. T. M., van Rosmalen, F., Koeze, J., Keus, F., van Kuijk, S. M. J., Castela Forte, J., Schnabel, R. M., Driessen, R. G. H., van Herpt, T. T. W., Sels, J.-W. E. M., Bergmans, D. C. J. J., Lexis, C. P. H., van Doorn, W. P. T. M., Meex, S. J. R., Xu, M., Borrat, X., Cavill, R., van der Horst, I. C. C., & van Bussel, B. C. T. (2024). Deep embedded clustering generalisability and adaptation for integrating mixed datatypes: Two critical care cohorts. *Scientific Reports*, *14*(1), Article 1045. https://doi.org/10.1038/s41598-024-51699-z

Formolli, M., Kleiven, T., & Lobaccaro, G. (2023). Assessing solar energy accessibility at high latitudes: A systematic review of urban spatial domains, metrics, and parameters. *Renewable and Sustainable Energy Reviews*, *177*, Article 113231. https://doi.org/10.1016/j.rser.2023.113231

Hao, H., & Wang, Y. (2024). *A deep learning representation of spatial interaction model for resilient spatial planning of community business clusters*. arXiv.

Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., & Zipf, A. (2023). A spatio-temporal analysis investigating completeness and inequalities of global urban building data in OpenStreetMap. *Nature Communications*, *14*(1), Article 3985. https://doi.org/10.1038/s41467-023-39698-6

Hu, Z., Zhang, L., Shen, Q., Chen, X., Wang, W., & Li, K. (2023). An integrated framework for residential layout designs: Combining parametric modeling, neural networks, and multi-objective optimization for outdoor activity space optimization. *Alexandria Engineering Journal*, *80*, 202–216. https://doi.org/10.1016/j.aej.2023.08.049

Kalliomäki, H., Oinas, P., & Salo, T. (2024). Innovation districts as strategic urban projects: The emergence of strategic spatial planning for urban innovation. *European Planning Studies*, *32*(1), 78–96. https://doi.org/10.1080/09654313.2023.2216727

Ko, J., Ennemoser, B., Yoo, W., Yan, W., & Clayton, M. J. (2023). Architectural spatial layout planning using artificial intelligence. *Automation in Construction*, *154*, Article 105019. https://doi.org/10.1016/j.autcon.2023.105019

Lan, H., Gou, Z., & Hou, C. (2022). Understanding the relationship between urban morphology and solar potential in mixed-use neighborhoods using machine learning algorithms. *Sustainable Cities and Society*, *87*, Article 104225. https://doi.org/10.1016/j.scs.2022.104225

Li, J., & Li, C. (2024). Characterizing urban spatial structure through built form typologies: A new framework using clustering ensembles. *Land Use Policy*, *141*, Article 107166. https://doi.org/10.1016/j.landusepol.2024.107166

Lu, Y., Chen, Q., Yu, M., Wu, Z., Huang, C., Fu, J., Yu, Z., & Yao, J. (2023). Exploring spatial and environmental heterogeneity affecting energy consumption in commercial buildings using machine learning. *Sustainable Cities and Society*, *95*, Article 104586. https://doi.org/10.1016/j.scs.2023.104586

Montero, G., Caruso, G., Hilal, M., & Thomas, I. (2023). A partition-free spatial clustering that preserves topology: Application to built-up density. *Journal of Geographical Systems*, *25*(1), 5–35. https://doi.org/10.1007/s10109-022-00396-4

Shen, T., Wu, J., Yuan, S., Kong, F., & Liu, Y. (2024). Analysis of urban spatial morphology in Harbin: A study based on building characteristics and driving factors. *Sustainability*, *16*(20), Article 9072. https://doi.org/10.3390/su16209072

Sun, D., Shen, T., Yang, X., Huo, L., & Kong, F. (2024). Research on a multi-scale clustering method for buildings taking into account visual cognition. *Buildings*, *14*(10), Article 3310. https://doi.org/10.3390/buildings14103310

Usui, H. (2024). Relative spatial variability in building heights and its spatial association: Application for the spatial clustering of harmonious and inharmonious building heights in Tokyo. *Environment and Planning B*: *Urban Analytics and City Science*, *51*(4), 987–1002. https://doi.org/10.1177/23998083231204691

Vera, C., Lucchini, F., Bro, N., Mendoza, M., Löbel, H., Gutiérrez, F., Dimter, J., Cuchacovic, G., Reyes, A., Valdivieso, H., Alvarado, N., & Toro, S. (2022). Learning to cluster urban areas: Two competitive approaches and an empirical validation. *EPJ Data Science*, *11*(1), Article 62. https://doi.org/10.1140/epjds/s13688-022-00374-2

Wang, J., Hu, Y., & Duolihong, W. (2023). Diagnosis and planning strategies for quality of urban street space based on street view images. *ISPRS International Journal of Geo-Information*, *12*(1), Article 15. https://doi.org/10.3390/ijgi12010015

Wu, P., Zhang, Z., Peng, X., & Wang, R. (2024). Deep learning solutions for smart city challenges in urban development. *Scientific Reports*, *14*(1), Article 5176. https://doi.org/10.1038/s41598-024-55928-3

Yang, C., Liu, T., & Zhang, S. (2022). Using flickr data to understand image of urban public spaces with a deep learning model: A case study of the haihe river in Tianjin. *ISPRS International Journal of Geo-Information*, *11*(10), Article 497. https://doi.org/10.3390/ijgi11100497

Yang, D., Zhao, J., & Xu, P. (2024). Deep learning-based approach for optimizing urban commercial space expansion using artificial neural networks. *Applied Sciences*, *14*(9), Article 3845. https://doi.org/10.3390/app14093845

Zhang, D., Kong, Q., & Shen, M. (2023). Does polycentric spatial structure narrow the urban-rural income gap? – Evidence from six urban clusters in China. *China Economic Review*, *80*, Article 101999. https://doi.org/10.1016/j.chieco.2023.101999

Zheng, Y., Lin, Y., Zhao, L., Wu, T., Jin, D., & Li, Y. (2023). Spatial planning of urban communities via deep reinforcement learning. *Nature Computational Science*, *3*(9), 748–762. https://doi.org/10.1038/s43588-023-00503-5