# Reimagining City Configuration: Automated Urban Planning via Adversarial Learning

Dongjie Wang
wangdongjie@Knights.ucf.edu
University of Central Florida
Orlando, Florida, United States

Yanjie Fu*
yanjie.fu@ucf.edu
University of Central Florida
Orlando, Florida, United States

Pengyang Wang
pengyang.wang@knights.ucf.edu
University of Central Florida
Orlando, Florida, United States

Bo Huang
bohuang@cuhk.edu.hk
The Chinese University of Hong Kong
Hong Kong, China

Chang-Tien Lu
ctlu@vt.edu
Virginia Tech
Virginia, United States

## ABSTRACT

Urban planning refers to the efforts of designing land-use configurations. Effective urban planning can help to mitigate the operational and social vulnerability of a urban system, such as high tax, crimes, traffic congestion and accidents, pollution, depression, and anxiety. Due to the high complexity of urban systems, such tasks are mostly completed by professional planners. But, human planners take longer time. The recent advance of deep learning motivates us to ask: can machines learn at a human capability to automatically and quickly calculate land-use configuration, so human planners can finally adjust machine-generated plans for specific needs? To this end, we formulate the automated urban planning problem into a task of learning to configure land-uses, given the surrounding spatial contexts. To set up the task, we define a land-use configuration as a longitude-latitude-channel tensor, where each channel is a category of POIs and the value of an entry is the number of POIs. The objective is then to propose an adversarial learning framework that can automatically generate such tensor for an unplanned area. In particular, we first characterize the contexts of surrounding areas of an unplanned area by learning representations from spatial graphs using geographic and human mobility data. Second, we combine each unplanned area and its surrounding context representation as a tuple, and categorize all the tuples into positive (well-planned areas) and negative samples (poorly-planned areas). Third, we develop an adversarial land-use configuration approach, where the surrounding context representation is fed into a generator to generate a land-use configuration, and a discriminator learns to distinguish among positive and negative samples. Finally, we devise two new measurements to evaluate the quality of land-use configurations and present extensive experiment and visualization results to demonstrate the effectiveness of our method.

*Concat Author.


*SIGSPATIAL '20, November 3–6, 2020, Seattle, WA, USA*

ACM ISBN 978-1-4503-8019-5/20/11…$15.00
https://doi.org/10.1145/3397536.3422268

## CCS CONCEPTS

• **Social and professional topics** → *Professional topics*; *Computing and business*; Automation;

## KEYWORDS

urban planning, representation learning, generative adversarial networks, graph neural networks

**ACM Reference Format:**
Dongjie Wang, Yanjie Fu, Pengyang Wang, Bo Huang, and Chang-Tien Lu. 2020. Reimagining City Configuration: Automated Urban Planning via Adversarial Learning . In *28th International Conference on Advances in Geographic Information Systems (SIGSPATIAL '20), November 3–6, 2020, Seattle, WA, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3397536.3422268

## 1 INTRODUCTION

Urban planning is an interdisciplinary and complex process that involves with public policy, social science, engineering, architecture, landscape, and other related field. In this paper, we refer urban planning to the efforts of designing land-use configurations, which is the reduced yet essential task of urban planning. Effective urban planning can help to mitigate the operational and social vulnerability of a urban system, such as high tax, crimes, traffic congestion and accidents, pollution, depression, and anxiety.

Due to the high complexity of urban systems, such planning tasks are mostly completed by professional planners. But, human planners take longer time. The recent advance of deep learning, particularly deep adversarial learning, provide a great potential for teaching a machine to imagine and create. This observation motivates us to rethink urban planning in the era of artificial intelligence: What roles does deep learning play in urban planning? Can machines develop and learn at a human capability to automatically and quickly calculate land-use configurations? In this way, machines can be planning assistants and human planners can finally adjust machine-generated plans for specific needs.

All of the above evidence has shown that it is appealing to develop a data-driven AI-enabled automated urban planner. However, three unique challenges arise to achieve the goal: (1) How can we quantify a land-use configuration plan? (2) How can we develop a machine learning framework that can learn the good and the bad of existing urban communities in terms of land-use configuration

policies? (3) How can we evaluate the quality of generated land-use configurations? Next, we will introduce our research insights and solutions to address the challenges.

First, as we aim to teach a machine to reimagine the land-use configuration of an area, it is naturally critical to define a machine-perceivable structure for a land-use configuration. In practice, the land-use configuration plan of a given area is visually defined by a set of Point of Interests (POIs) and their corresponding locations (e.g., latitudes and longitudes) and urban functionality categories (e.g., shopping, banks, education, entertainment, residential). A close look into such visually-perceived land-use configuration can reveal that the land-use configuration is indeed a high-dimensional indicator that illustrates what and where we should put into an unplanned area. There is not just location-location statistical autocorrelation but also location-functionality statistical autocorrelation in a land-use configuration. To capture such statistical correlations, we propose to represent a land use configuration plan as a latitude-longitude-channel tensor, where each channel is a specific category of POIs that are distributed across the unplanned area, and the value of an entry in the tensor is the number of POIs. In this way, the tensor can describe not just the location-location interaction of POIs, but also location-function interaction of POIs.

Second, after we define the quantitative version of a land-use configuration, our second question is that how we can teach a machine to automatically generate a land-use configuration? We analyze large-scale urban residential community data, and identify an important observation: 1) an urban community can be viewed as an attributed node in a socioeconomic network (city), and this node proactively interacts with surrounding nodes (environments); 2) the coupling, interaction, and coordination of a community and surrounding environments significantly influence the livability, vibrancy, and quality of a community. Based on this observation, we propose to convert the land-use configuration planning problem as a new objective: to teach a machine to generate a land-use configuration tensor given the surrounding context/area. In other words, the problem is reduced into the objective of learning a conditional probability function that maps a surrounding context representation to a well-planned configuration tensor, instead of a poorly-planned configuration tensor. The recently emerging deep adversarial learning provides a great potential to address the reduced objective. We reformulate the task into an adversarial learning paradigm, in which: 1) A neural generator is analogized as a machine planner that generates a land-use configuration; 2) The generator generates a configuration in terms of a pattern feature representation of surrounding spatial contexts; 3) The surrounding context representation is learned via self-supervised representation learning collectively from spatial graphs. 4) A neural discriminator is to classify whether the generated land-use configuration is well-planned (positive) or poorly-planned (negative). 5) A new mini-max loss function is constructed to guide the generator to learn from the goods of well-planned areas and the bads of poorly-planned areas.

Third, it is traditionally a very challenging open question to evaluate the quality of a generated land-use configuration. The most solid and sound validation is to collaborate with urban developer and city governments to implement an AI-generated configuration into an unplanned area, and observe the development of the area in the following years. However, this is not realistic. In this paper, we exploit two strategies to assess the generated configurations: 1) We build one scoring model that outputs the score of a configuration by learning from training data. Specifically, we utilize a machine learning model to learn the data distribution of the original land-use reconfiguration samples. Thus, after we obtain the generated solutions, the scoring model has the ability to give a score. 2) We invite experienced regional planning experts to evaluate the quality of the generated solutions. In this paper, we let human experts to perform analysis on multiple case studies.

In summary, in this paper, we develop an adversarial learning framework to generate effective land-use configurations by learning from urban geography, human mobility, and socioeconomic data. Specifically, our contributions are: 1) We develop a latitude-longitude-channel tensor to quantify a land-use configuration plan. 2) We propose a socioeconomic interaction perspective to understand urban planning as a process of optimizing the coupling between a community and surrounding environments. 3) We reformulate the automated urban planning problem into an adversarial learning framework that maps surrounding spatial contexts into a configuration tensor by a machine generator. 4) Although evaluation is challenging, we develop multiple aspects to conduct extensive experiment and visualization with real-world data to show the value of our method.

## 2 PROBLEM STATEMENT AND FRAMEWORK OVERVIEW

### 2.1 Definitions

*2.1.1 Central area and its Contexts.* In this paper, a central area is a square area that is centered on a geographical location (i.e., latitude and longitude), where is an unplanned area. In this study, the area of a central area is $1km^2$. The contexts of a central area wrap the residential community from different directions.


**Figure 1: The geographical definitions of a central area and its surrounding spatial contexts.**

Figure 1 represents the spatial relationship between a central area and its contexts. In addition, we find that there are many POIs in the central area and its contexts. Intuitively, the future development of an unplanned area is affected by its contexts.

## 2.2 Problem Statement

Urban planning is a complex field. An excellent urban planning solution requires experts who own the amount of specific urban planning knowledge and experiences to spend much time designing. In order to reduce the heavy burden of the experts and generate suitable urban planning solutions objectively, we propose an automatic urban planner that produces excellent solutions based on the context environments. Here, we simplify the meaning of urban planning into land-use configuration, which makes the question easier to model. Formally, assuming a virgin area is $R$, the contexts of $R$ are $[C_1 \sim C_8]$, and the land-use configuration plan is $\mathrm{M}$ that structure is a multi-channel image, one channel represents one POI category data distribution. Given the explicit feature $\mathrm{F}$ that describes the situation of the context environments, where the matrix $\mathrm{F} \in \mathbb{R}^{8\times K}$, 8 is the number of contexts and $K$ is the dimension of the explicit feature vector of each context.

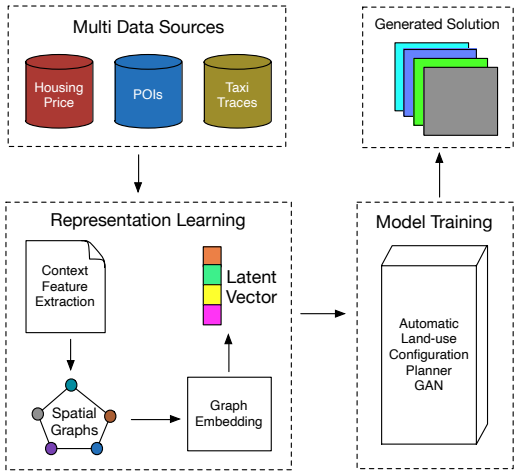



**Figure 2: An overview of the proposed framework. The proposed framework includes four steps: 1) we first collect multiple data sources such as urban community related data (housing prices), point of interests data, and human mobility data (taxicab GPS traces). We then propose a spatial graph representation learning to learn the representation of surrounding contexts. Later, we develop an adversarial land use configuration machine to automate planning and generate recommended configurations.**

The purpose of our framework is to take the explicit feature vector $\mathrm{F}$ as input and output the corresponding excellent land-use configuration solution $\mathrm{M}$.

## 2.3 Framework Overview

Figure 2 shows an overview of our proposed method (LUCGAN). This framework has two main parts: (i) learning representation of the contexts of the virgin area; (ii) generating an excellent land-use configuration solution for the virgin area. In the first part, we first extract explicit features of the contexts from value-added space, POI distribution, public and private transportation conditions. Then, we construct a graph structure to capture the geographical spatial relationship between the virgin area and its contexts. Afterward, we map the explicit features of contexts to the graph as attributes of corresponding nodes. The attributed spatial graph incorporates all characteristics of contexts together. Next, we utilize variational graph auto-encoder (VGAE) to obtain the latent representation of the contexts. Thus, we get the final representation of the contexts of virgin areas through the first part. In the second part, we input the latent representation of the contexts, excellent land-use configuration samples, and terrible land-use configuration samples into an extended GAN. The extended GAN is capable to generate the land-use configuration solution based on the contexts embedding. Moreover, we customize a new GAN loss which makes the model learns the distribution of excellent plans and keeps away from the terrible plans. Finally, when the model converges, the generator of the extended GAN can produce suitable and excellent land-use configuration solutions in an objective angle based on the latent context embedding.

# 3 AUTOMATIC PLANNER FOR LAND-USE CONFIGURATION

In this section, we first introduce how to quantify surrounding context/area. Then, we detail the measurement of how to evaluate the quality of land-use configuration solutions. Finally, we describe how to train a generative model to learn an automated urban planner.

## 3.1 Explicit Feature Extraction for Context Environments

The land-use configuration solution of an unplanned area has a strong relationship with its contexts. For example, if there are many commercial zones in the contexts of the unplanned area, we should avoid the redundancy of the same category POI in the planning. This is because we can make the unplanned area owns different function compared with its contexts, which is beneficial for the development and communication among the virgin area and its contexts. Thus we grab the intrinsic characteristics of the contexts completely by extracting multiple explicit features.

There are lots of indicators to describe context environments. Here, we select four views to capture the features of the contexts:

(1) **Value-added Space.** In common, the variation of house price reflects the value-added space of one area. Thus, we calculate the changing trend of house price of the contexts $[C_1 \sim C_8]$ in continued six months. Here, we take the context $C_1$ as an example to explain the calculation process. First, We obtain the housing price list among $t$ months. Then, we calculate the changing trend of house price by using the current house price value subtract the previous house price value. So we get the changing trend of $C_1$ as $\mathrm{v}_1 = [v_1^1, v_1^2, ..., v_1^{t-1}]$, where $v_1^i$ represents the value of the changing trend at i-th month. Finally, we collect the house price changing trend of all contexts together. The collected result is denoted as $\mathrm{V} = [\mathrm{v}_1, \mathrm{v}_2, ..., \mathrm{v}_8]$, where the matrix $\mathrm{V} \in \mathbb{R}^{8\times t-1}$.

(2) **POI Ratio.** Since various POIs provide diverse services to residents, the ratio of different types of POIs is a good indicator for indicating the functions of the area. Therefore, we calculate the POI ratio of the contexts $[C_1 \sim C_8]$. Here, we take $C_1$ as an example to explain the calculation process. First, we sum up the count of each POI category to form a feature vector. Then, we divide each item in the feature

vector by the sum of all POI categories. We obtain the POI ratio of $C_1$, denoted by $\mathrm{r}_1 = [r_1^1, r_1^2, ..., r_1^m]$, where $r_1^i$ represents the ratio of i-th POI category in $C_1$ and $m$ is the total number of POI categories. Finally, we collect the POI ratio of all contexts together. The collected result is denoted as $\mathrm{R} = [\mathrm{r}_1, \mathrm{r}_2, ..., \mathrm{r}_8]$, where the matrix $\mathrm{R} \in \mathbb{R}^{8\times m}$, $m$ is the number of POI categories.

(3) **Public Transportation.** Public transportation is one popular travel mode due to its convenience and cheapness. So public transportation is a vital factor to be considered to describe the human mobility patterns. Thus, we extract features that related to public transportation to describe the public traffic situation of the contexts $C_1 \sim C_8$. We take $C_1$ as an example to show the calculation details. We calculate the feature vector of public transportation from five perspectives: (1) the leaving volume of $C_1$ in one day, denoted by $\mathrm{o}_1^1$; (2) the arriving volume of $C_1$ in one day, denoted by $\mathrm{o}_1^2$; (3) the transition volume of $C_1$ in one day, denoted by $\mathrm{o}_1^3$; (4) the density of bus stop of $C_1$, denoted by $\mathrm{o}_1^4$, which reflects the number of bus stop in per square meter; (5) the average balance of smart card of $C_1$, denoted by $\mathrm{o}_1^5$, which shows the economic expenditure of people in the travel field. The public transportation feature vector of $C_1$ can be denoted as $[o_1^1, o_1^2, ..., o_1^5]$. Finally, we collect the public transportation feature vectors of all contexts together. The collected result is denoted as $\mathrm{O} = [\mathrm{o}_1, \mathrm{o}_2, ..., \mathrm{o}_8]$, where the matrix $\mathrm{O} \in \mathbb{R}^{8\times 5}$.

(4) **Private Transportation.** Taxi is another important tool for people traveling. The taxi trajectory data reflects the people's flow count and the traffic congestion situation of an area. Thus, we explore the features of the private transportation condition of the contexts $[C_1 \sim C_8]$. Here, we take $C_1$ as an example to illustrate the calculation process. We count the features of private transportation from the following 5 perspectives: (1) the leaving volume of $C_1$ in one day, denoted by $u_1^1$; (2) the arriving volume of $C_1$ in one day, denoted by $u_1^2$; (3) the transition volume of $C_1$ in one day, denoted by $u_1^3$; (4) in $C_1$, the average driving velocity of a taxi in one hour, denoted by $u_1^4$; (5) in $C_1$, the average commute distance for a taxi, denoted by $u_1^5$; Then, the feature vector of private transportation can be denoted as $[u_1^1, u_1^2, ..., u_1^5]$. Ultimately, we collect the private transportation feature vectors of all contexts together. The collected result is denoted as $\mathrm{U} = [\mathrm{u}_1, \mathrm{u}_2, ..., \mathrm{u}_8]$, where the matrix $\mathrm{U} \in \mathbb{R}^{8\times 5}$.

After that, we obtain the explicit feature set of the contexts $C_1 \sim C_8$. The set contains four kinds of features $[\mathrm{V}, \mathrm{R}, \mathrm{O}, \mathrm{U}]$, which describe the context environments from four perspectives.

## 3.2 Explicit Features as Node Attributes: Constructing the Spatial Attributed Graph

The context environments wrap the residential community from different directions, resulting in spatial correlation among areas. Such phenomenon motivates us to exploit spatial graphs to capture such spatial correlations. Specifically, Figure 3 shows the graph structure, where the blue nodes represent the contexts; the orange node is the residential community; the edge between two nodes reflects the connectivity between them.

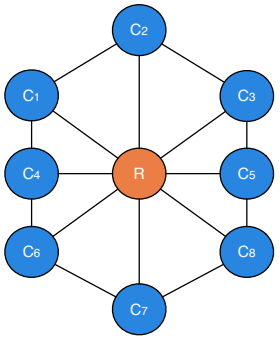



**Figure 3: The graph structure between residential community and its surrounding spatial contexts.**

Then, in order to fuse the spatial relationship and explicit features of the contexts, we construct a spatial attributed graph structure. Formally, we map the explicit features to the spatial graph based on the corresponding context node as the node attribute. Figure 4 expresses the construction process of the spatial attributed graph. This graph not only contains the explicit feature vector of the contexts but also includes the spatial relations among them.

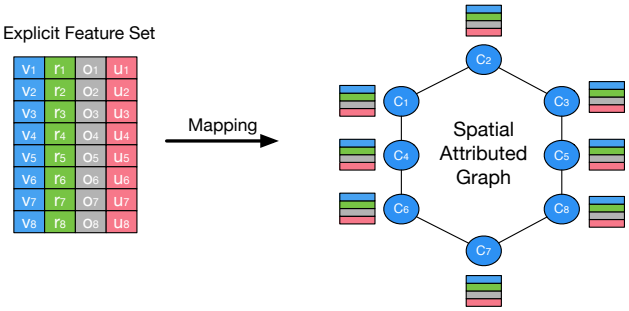



**Figure 4: The illustration of constructing spatial attributed graphs: Each feature vector is mapped to the corresponding nodes by a column-wise strategy.**

## 3.3 Learning Representation of the Spatial Attributed Graph

Figure 5 shows that we develop a spatial representation learning framework to preserve and fuse the explicit features and spatial relationship in the contexts. Formally, we denote the spatial attributed graph $G$ by $G = (\mathrm{X}, \mathrm{A})$, where A is the adjacency matrix that expresses the accessibility among different nodes; $X$ is the feature matrix of the graph, here, $\mathrm{X} = [\mathrm{V}, \mathrm{R}, \mathrm{O}, \mathrm{U}]$, and the concatenation direction is row-wise. In order to get the latent graph embedding z, we minimize the reconstruction loss between original graph $G$ and the reconstructed graph $\widehat{G}$ by an encoding-decoding framework.

The encoder part owns two Graph Convolutional Network (GCN) layers. The first GCN layer takes X and A as input and output the feature matrix of low-dimensional space $\widehat{\mathrm{X}}$. Thus, the encoding process can be formulated as:

$$\widehat{\mathrm{X}} = GCN_1(\mathrm{X}, \mathrm{A}) = RELU(\widehat{\mathrm{D}}^{-\frac{1}{2}}\mathrm{A}\widehat{\mathrm{D}}^{-\frac{1}{2}}\mathrm{X}\mathrm{W}_1) \quad (1)$$

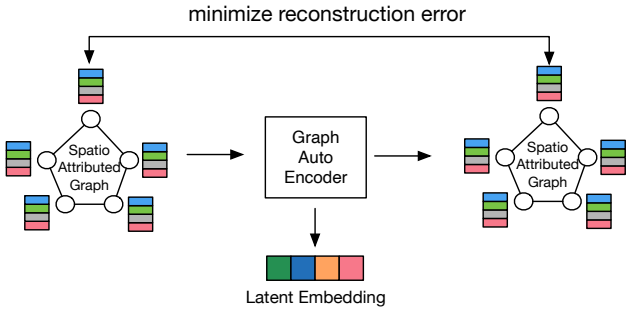



**Figure 5: The proposed representation learning model to obtain surrounding context representations by minimizing the reconstruction loss of spatial attributed graphs.**

where $\widehat{\mathrm{D}}$ is the diagonal degree matrix, $\mathrm{W}_1$ is the weight matrix of the $GCN_1$, and the whole layer is activated by *RELU* function.

Owing to the latent embedding z is sampled from a prior Normal Distribution, the second GCN layer is responsible for assessing the parameters of the prior distribution. Formally, the second GCN layer takes $\widehat{\mathrm{X}}$ and A as input and then outputs the mean value $\boldsymbol{\mu}$ and the variance value $\boldsymbol{\delta}^2$. So the calcuation process of the second GCN layer can be formulated as:

$$\boldsymbol{\mu}, log(\boldsymbol{\delta}^2) = GCN_2(\mathrm{X}, \mathrm{A}) = \widehat{\mathrm{D}}^{-\frac{1}{2}}\mathrm{A}\widehat{\mathrm{D}}^{-\frac{1}{2}}\widehat{\mathrm{X}}\mathrm{W}_2 \tag{2}$$

where $\mathrm{W}_2$ is the weight matrix of $GCN_2$. Next, we use the reparameterization trick to approximate the sample operation to obtain the latent representation z:

$$\mathrm{z} = \boldsymbol{\mu} + \boldsymbol{\delta} \times \epsilon \tag{3}$$

where $\epsilon \sim \mathcal{N}(0, 1)$.

The decoding module takes the z as input and then outputs the reconstructed adjacent matrix $\widehat{\mathrm{A}}$. So the decoding step can be formulated as:

$$\widehat{\mathrm{A}} = \sigma(\mathrm{z}\mathrm{z}^T) \tag{4}$$

where $\sigma$ represents the decoding layer is activated by sigmoid function. Moreover, $\mathrm{z}\mathrm{z}^T$ can be converted to $\|z\| \left\|\mathrm{z}^T\right\| \cos\theta$. The inner product operation is beneficial capture the spatial correlation among different contexts.

During the training phase, we minimize the joint loss function $\mathcal{L}$ that is denoted as:

$$\mathcal{L} = \sum_{i=1}^{N} \underbrace{KL[q(\mathrm{z}|\mathrm{X}, \mathrm{A})||p(\mathrm{z})]}_{\text{KL Divergence between } q(.) \text{ and } p(.)} + \overbrace{\sum_{j=1}^{S} \left\|\mathrm{A} - \widehat{\mathrm{A}}\right\|^2}^{\text{Loss between A and } \widehat{\mathrm{A}}} \tag{5}$$

where $N$ is the dimension of z; $S$ is the total number of the nodes in A; $q$ represents the real distribution of z; $p$ represents the prior distribution of z. $\mathcal{L}$ includes two parts, the first part is the Kullback-Leibler divergence between the standard prior distribution $\mathcal{N}(0, 1)$ and the distribution of z, and the second part is the squared error between A and $\widehat{\mathrm{A}}$. The training process try to make the $\widehat{\mathrm{A}}$ close to A and let the the distribution of z similar to $\mathcal{N}(0, 1)$.

Finally, we utilize global average aggregation for z to get the graph level representation, which is the latent representation of all context environments.

## 3.4 Land-use Configuration Quantification and the Quality Measurement

Land-use configuration indicates the location of different types of POIs, which expects an appropriate format of quantification for accommodating with a learning model. To that end, we regard the POI distribution of one area as the land-use configuration And then, we construct a multi-channel tensor to represent the land-use configuration, where each channel is the POI distribution across the geospatial area corresponding to one POI category. Figure 6 shows an example of a land-use configuration. We first divide an unplanned area into $n \times n$ squares, then we sum up the number of each POI category in each square respectively. Here, one POI category constructs one channel of the land-use configuration solution. We obtain a land-use configuration as a multi-channel tensor.

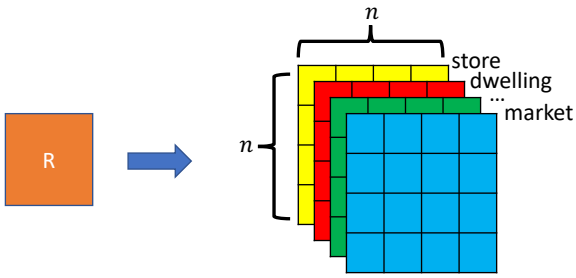



**Figure 6: The construction of longitude, latitude, channel configuration tensor where the value of each entry is the number of POIs with respect to a specific category in a specific latitude range and a specific longitude range.**

Next, the other big challenge is how to evaluate the quality of land use configuration of the residential community? Because the urban planning is a complex field, urban planning specialists always evaluate the quality of land-use configuration solution from multiple aspects. In our framework, we provide a quality hyperparameter $Q$ for users, they can set the value of $Q$ to distinguish the quality of land-use configuration solution. In our experiment, we choose the POI diversity and the check-in frequency of an area as the quality standard. Formally, we first calculate the total number of mobile check-in events of an area, denoted by $freq$, and the diversity of POI of an area, denoted by $div$. We then incorporate the two indicators into together by $Q = \frac{2 \times freq \times div}{freq + div}$ [20]. If $Q > 0.5$, the solution is regarded as an excellent solution. Otherwise, it is justified as a terrible solution.

## 3.5 Generating Excellent Land-use Configuration Solution by GAN

Generative adversarial networks (GAN) is a popular deep generative model. The framework of this technique is suitable to generate realistic data samples via an adversarial way. In the computer vision field, GAN achieves tremendous achievements. So, here, we utilize the GAN framework to generate excellent land-use configuration solutions of an unplanned area according to the representation of the context environments.

Figure 7 represents the structure of our automated land-use configuration planner. In common, the real land-use configuration includes two categories: excellent and terrible. The purpose of the

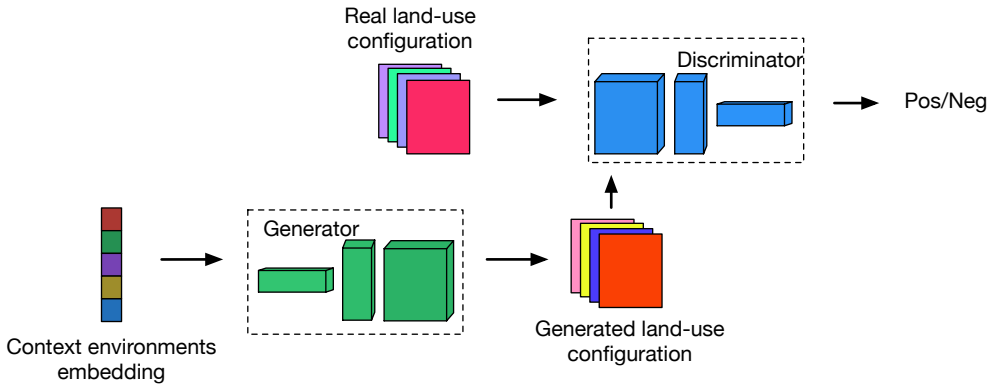



**Figure 7: Automatic land-use configuration planner**

automated planner is to generate the excellent land-use configuration plan based on the context embedding. Formally, we input the context embedding into the generator to generate the land-use configuration solution. In order to improve the generative ability, the discriminator classify the excellent plans as positive and the terrible plans as negative. Algorithm 1 shows the detail information about the training phase.

**Algorithm 1:** Minibatch adaptive moment estimation training of automatic land-use configuration model. We adjust one hyperparameter $f$ to change the update frequencies of the weight of the discriminator.

```
// start training.
for number of training iterations do
    // update discriminator firstly.
    for n steps do
        Sample minibatch of m excellent land-use
          configuration samples {E^1, E^2, ..., E^m}.
        Sample minibatch of m context information
          embedding samples {z^1, z^2, ..., z^m}.
        Generate land-use configuration samples by
          generator, {F^1, F^2, ..., F^m}. Here, F^i = G(z^i).
        Sample minibatch of m terrible land-use
          configuration samples {T^1, T^2, ..., T^m}.
        Update the discriminator by ascending its gradient:

        ∇_{θ_d} (1/m) Σ_{i=1}^{m} [log(D(E^i)) + log(D(1 − F^i))
        +log(D(1 − T^i))].
    // update generator secondly.
    Sample minibatch of m context information embedding
      samples {z^1, z^2, ..., z^m}.
    Update the generator by descending its gradient:

    ∇_{θ_g} (1/m) Σ_{i=1}^{m} log(1 − D(G(z^i))).
```

In algorithm 1, we first update the parameters of the discriminator module and fix the parameters of the generator module. We then feed the excellent, terrible, and generated land-use configuration samples into the discriminator module. Next, the discriminator outputs the classification result that is activated by the sigmoid function. It gives higher classification scores for excellent samples than terrible and generated samples. Next, we fix the discriminator module and update the parameter of the generator module. The contexts embedding vectors are feed into the generator, and output generated land-use configuration solutions. Afterward, the generated solutions are feed into the discriminator to justify the quality of them. We update the parameter of the generator module to improve the generated ability of itself. The update gradient comes from the justification result of the discriminator module. Finally, We obtain one automatic land-use configuration planner when the GAN model converges. If we obtain the context embedding, the discriminator can generate one excellent land-use configuration for the unplanned area.

## 4 EXPERIMENT RESULTS

In this section, we conduct extensive experiments and case studies to answer the following questions:

(1) Does our proposed automatic planner (LUCGAN) outperform the baseline methods?
(2) What is the difference between the context of excellent land-use configuration plans and terrible plans?
(3) How does generated land-use configuration plans look like?
(4) How many proportions of each POI category occupies in generated plans?
(5) What is the generated situations of different categories in the generated plan?

### 4.1 Data Description

We use the following datasets for our study:

(1) **Residential Community:** The residential community dataset contains 2990 residential communities in Beijing, where each residential community is associated with the information of its latitude and longitude.
(2) **POI:** The Beijing POI dataset includes 328668 POI records from 2011, where each POI item includes its latitude, longitude and the category. The POI information is shown in Table 1.

(3) **Taxi Trajectories:** The taxi trajectories dataset are collected from a Beijing taxi company, where each record contains trip ID, distance(m), travel time(s), average speed(km/h), pick-up and drop off time, pick-up and drop-off point.
(4) **Public Transportation:** This dataset logs the transactions of buses in Beijing between 2012 and 2013. After analyzing the dataset, it contains 1734247 bus trips, 718 bus lines. We use this dataset to obtain the public transportation situation.
(5) **House Price:** The house price dataset includes continuous five months house price data of each residential community in Beijing between 2011 and 2012, which is collected from Soufang website.
(6) **Check-In:** This dataset is the weibo check-in records in Beijing between 2011 and 2013, where each check-in item has its longitude, latitude, check-in time and check-in place. We utilize this dataset to analyze the vibrancy of one area.

**Table 1: POI category list**

| code | POI category | code | POI category |
|---|---|---|---|
| 0 | road | 10 | tourist attraction |
| 1 | car service | 11 | real estate |
| 2 | car repair | 12 | government place |
| 3 | motorbike service | 13 | education |
| 4 | food service | 14 | transportation |
| 5 | shopping | 15 | finance |
| 6 | daily life service | 16 | company |
| 7 | recreation service | 17 | road furniture |
| 8 | medical service | 18 | specific address |
| 9 | lodging | 19 | public service |

## 4.2 Evaluation Metrics

Because evaluating the quality of the urban land-use configuration is an open question, there is no standard measurement. In this paper, we evaluate the quality of generated planning solution from multiple aspects to express the effectiveness of our framework: (1)**Scoring Model**. We build a random forest model based on the excellent and terrible land-use configuration plans. The model is capable of giving higher scores for excellent land-use configuration plans and provide lower scores for terrible plans. When we get the generate land-use configuration solutions, the scoring model can be utilized to quantify the quality of the generated solutions. (2) **Visualization**. In order to explore the generated solutions, we select one representative sample to visualize from multiple aspects. We can observe the solutions directly in this way. It is helpful to learn the difference between our planner and other baselines .

## 4.3 Baseline Methods

We compare the performances of our framework(LUCGAN) against the following three baseline methods:

(1) **VAE**: is an encoder-decoder paradigm algorithm. The encoder encodes image data into latent embedding; the decoder decodes the embedding into the original data. In this experiment, we input excellent land-use configuration into VAE to learn the distribution of excellent solutions by minimizing the reconstruction loss. Then, we utilize the decoder to generate the solution based on the context environment embedding when VAE converges.
(2) **AVG**: generates the land-use configuration by calculating the mean value of all excellent land-use plans, which reflects the average level of all excellent samples. But this method can not provide a customized solution based on different context environments.

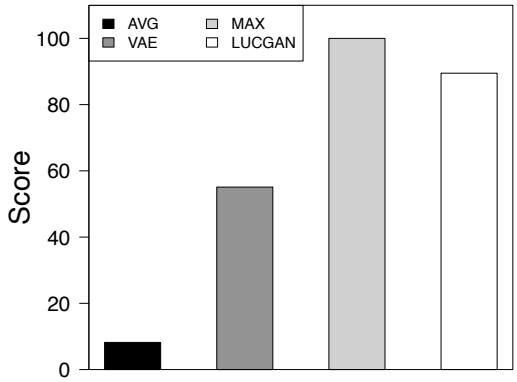



**Figure 8: The quality score for different generated methods.**

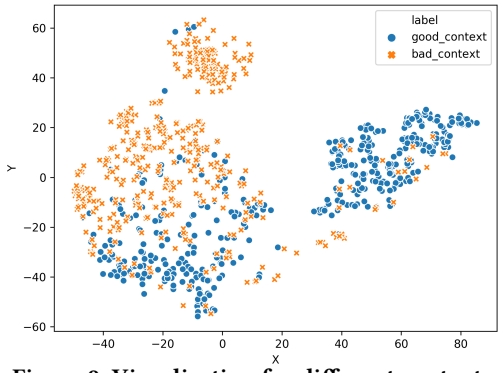


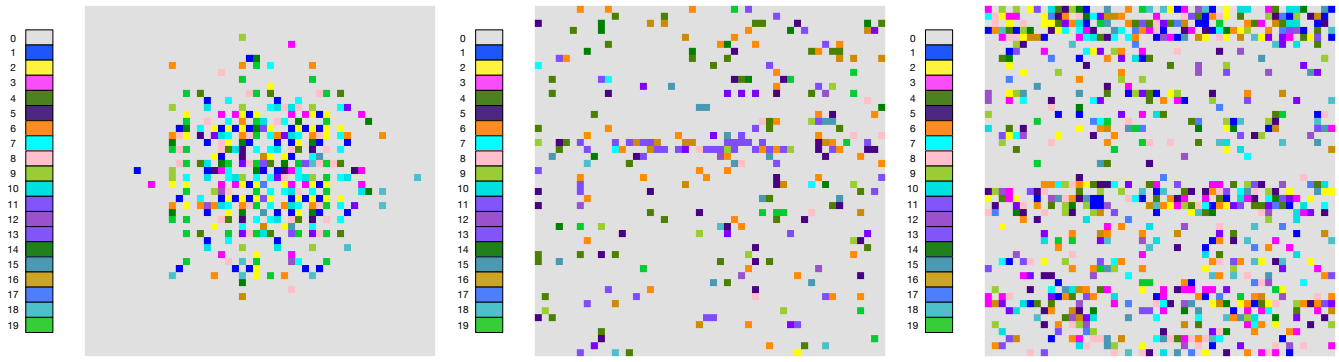

**Figure 9: Visualization for different contexts.**

(3) **MAX**: generates the land-use configuration solutions by applying max operation on all excellent land-use plans. The result of this method reflects the most dominated POI categories in each geographical block. The same to AVG, MAX also can not generate a customized solution based on different context environments.

We conduct all experiments on a x64 machine with Intel i9-9920X 3.50GHz CPU, 128GB RAM and Ubuntu 18.04.

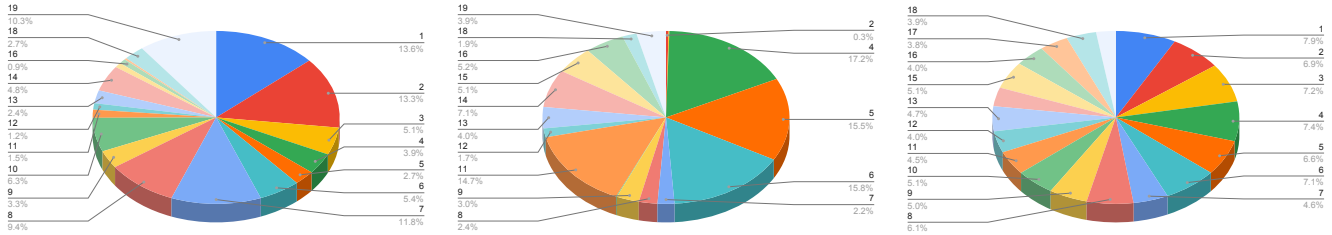
## 4.4 Overall Performance

Figure 8 shows the quality score produced by the scoring model for different generated methods. An interesting phenomenon is that the MAX method ranks the 1st compared with other methods. A possible explanation is the scoring model only captures the distribution of original excellent plans. The MAX method incorporates all excellent plans by max operation, so the generated solution reflects the dominated POI categories of each geographical block, which is also inherent in the original data distribution. Therefore, the scoring model gives the highest score for the MAX method. Although the MAX method ranks the 1st, it does not indicate the MAX method is better than LUCGAN. This is because the MAX method only produces one kind of planning solution no matter what the context environment is. But the LUCGAN can customize



(a) LUCGAN (b) VAE (c) MAX

**Figure 10: Comparison of different generated land-use configuration solutions by different generated methods.**



(a) LUCGAN (b) VAE (c) MAX

**Figure 11: Comparison of the proportion of each POI category in different generated solutions.**

the solutions based on different context environment embedding. In addition, the score of LUCGAN is also high, which indicates the LUCGAN capture the intrinsic rule of excellent plan distribution. So the LUCGAN is an effective and flexible generated method for generating land-use configuration.

## 4.5 Study of the Context Environment

Intuitively, the context environment is vital for generating the land-use configuration. Different context environments produce different kinds of land-use configuration solutions. In order to explore the difference between the context environments of excellent plans and bad plans, we utilize T-SNE algorithm on the context environment embedding. We randomly choose 500 good and bad context embedding respectively to visualize.

Figure 9 is the visualization result of different kinds of context environments. We find that the pattern of good contexts and the pattern of bad contexts are discriminative, which proves that it is reasonable to exploit the context environment for generating the land-use configuration solutions.

## 4.6 Study of the Geographical Distribution Generated by Different Approaches

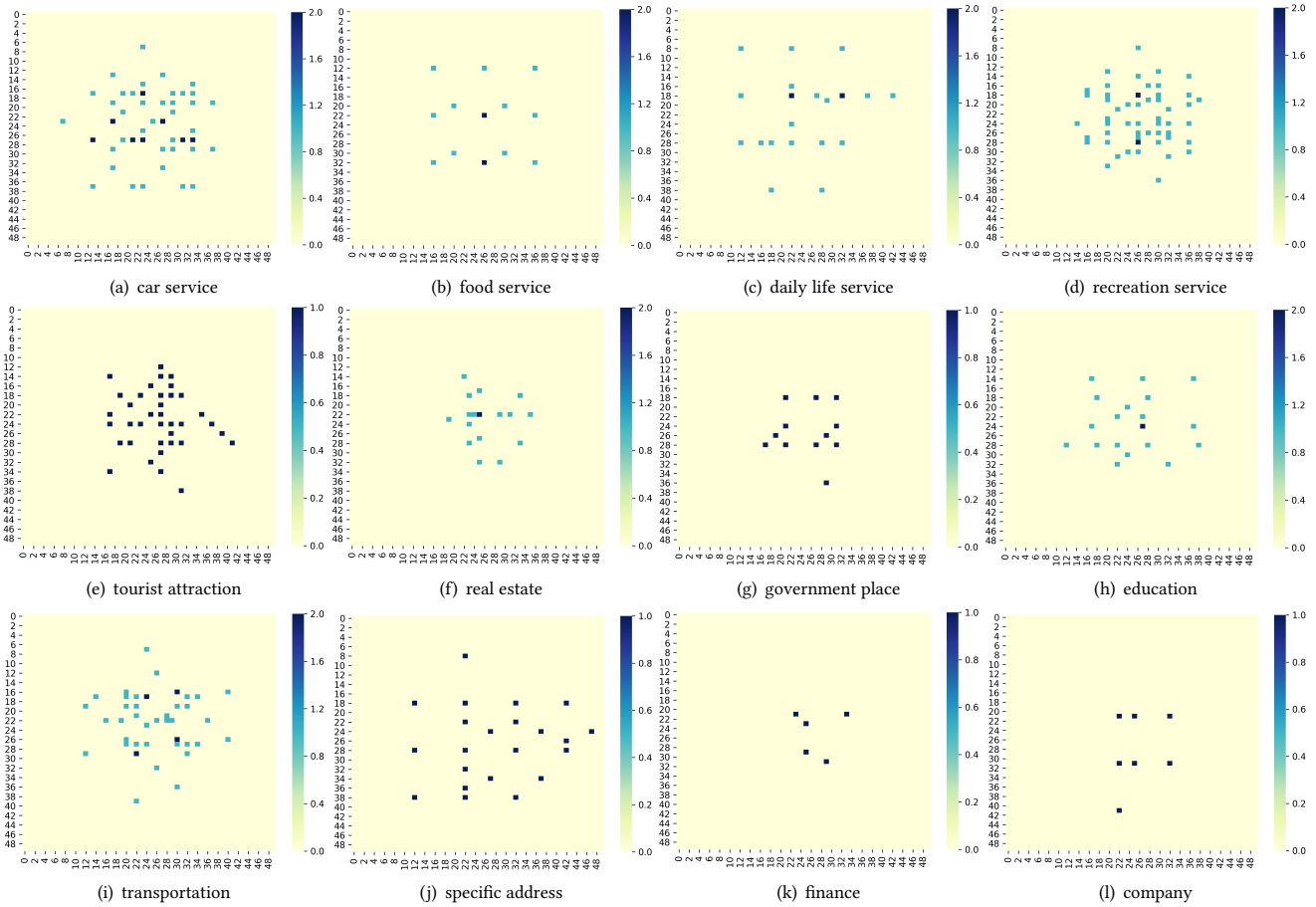
In order to observe the generated land-use configuration solution clearly, we pick up a representative solution to visualize. Owing to the generated solution has multiple channels and each channel owns lots of blocks, we merge these channels of the solution into one by we setting the dominated POI category as the final result for each geographical block. The merged solution reflects POI distribution in geographical spatial space.

Figure 10 shows the visualization result Different color blocks represent different POI categories. The generated solution by LUCGAN is regular and organized, different POI categories intersect with each other. But the distribution of generated solutions by VAE and MAX are so chaotic that there no clear promising patterns. This experiment indicates the better effectiveness of LUCGAN compared against other baseline methods.

## 4.7 Study of The POI Proportion Generated by Different Approaches

In this experiment, our purpose is to generate a vibrant residential community that owns diverse POI categories and lots of economic activities. After obtaining the generated solutions from different generated methods, we count the number of each POI category of different solutions respectively. Then we visualize the proportion of each POI category of the solutions.

Figure 11 shows the comparison of different generated methods. In Figure 11(a) , we observe that the LUCGAN generated result owns all POI categories, and 4 (food service), 5 (shopping), 6 (daily life service) and 7 (recreation service) occupy a big proportion in all POI categories. In addition, These four POI types are all related to the economic activities closely. Therefore, the LUCGAN generated solutions satisfy our design scheme. Figure 11(b) shows the POI proportion relation in VAE generated result. We find that 1 (car service) and 17 (road furniture) are missing, so the POI diversity of VAE is not complete. Figure 11(c) is the POI proportion relation in MAX generated result. Each POI category occupies balanced



(a) car service (b) food service (c) daily life service (d) recreation service

(e) tourist attraction (f) real estate (g) government place (h) education

(i) transportation (j) specific address (k) finance (l) company

**Figure 12: Visualization for different POI categories of one generated solution.**

proportion, which indicates the MAX method is stubborn. The flexibility of MAX is poorer than VAE and LUCGAN.

## 4.8 Study of The Generation for Each POI Category

We aim to examine the justifiability of the generated configuration for each POI category To that end, we visualize the POI distribution for each POI category. Due to the page limitation, we randomly select 12 POIs of a generated solution for visualization In Figure 12, the darker color block represents the larger POI number in the block. An interesting observation is that the POI distribution of different categories show their unique patterns. For example, transportation pots are more concentrated, while food service related POIs are more dispersed across the area; the distribution of car service spots is very similar to the recreation service, the possible reason is that recreation service spots may occupy many parking lots which potentially attract car services.

To sum up, through the above experiments, LUCGAN is capable to generate the land-use configuration solution based on the context environment embedding effectively and flexibly.

# 5 RELATED WORK

**Representation Learning.** The objective of representation learning is to obtain the low-dimensional representation of original data in latent space. In general, there are three types of representation learning models: (1) probabilistic graphical models; (2) manifold learning models; (3) auto-encoder models. The probabilistic graphical models build a complex Bayesian network system to learn the representation of uncertain knowledge buried in original data [14]. The manifold learning models infer low-dimensional manifold of original data based on neighborhood information by non-parametric approaches [24]. The auto-encoder models learn the latent representation by minimizing the reconstruction loss between original and reconstructed data [13]. In this paper, we utilize the auto-encoder paradigm to learn the representation of the spatial context environment. There are many successful spatial representation applications with the development of deep representation learning techniques [7, 8, 18, 21]. For instance, Wang et al. capture the feature of GPS trajectory data by utilizing spatio-temporal embedding learning skills, then the embedding is used to analyze driving behavior [19]. Du et al. propose a new spatial representation learning framework that captures the static and dynamic characteristics among the spatial entities. And they utilize the learned spatial representation to improve the performance of real estate price prediction. [6].

**Generative Adversarial Networks.** Generative Adversarial Networks is a hot research field in recent years [22, 23]. The GAN algorithms can be classified into three categories from task-driven

view . (1). semi-supervised learning GANs. Usually, a complete labeled data set is difficult to obtain, the semi-supervised learning GANs can utilize unlabeled data or partially labeled data to train an excellent classifier [5, 11]. For instance, Akcay et al. design a semi-supervised GAN anomaly detection framework that only uses normal data during the training phase [2]. (2). transfer learning GANs. Many researchers utilize the transfer learning GANs to transfer knowledge among different domains [10, 17]. For instance, Choi et al. build an unified generative adversarial networks to translate the images in different domains [4]. (3). reinforcement learning GANs. Generative models are incorporated with reinforcement learning (RL) to improve the model generative performance [16]. For instance, Ganin et al. combine reinforce learning and GAN to synthesize high-quality images [9].

**Urban Planning.** Urban planning is a complex research field [1]. The specialists need to consider lots of factors such as government policy, environmental protection to design suitable land-use configuration plan [12]. For example, partial researchers focus on constructing an urban planning solution for human health and well-being [3]. Simultaneously, many researchers form a land-use configuration solution based on the development of real estate [15]. Therefore, it is difficult to generate a suitable and excellent urban planning solution objectively. To the best of our knowledge, we first propose an automatic land-use configuration planner to solve this problem.

## 6 CONCLUSION REMARKS

In order to generate a suitable and excellent land-use configuration solution objectively and reduce the heavy burden of urban planning specialists, we proposed an automatic land-use configuration planner framework. This framework generates the land-use solution based on the context embedding of a virgin area. Specifically, we obtained the residential community and its context based on the latitude and longitude of residential areas firstly. we then extracted the explicit features of the context from three aspects: (1) value-added space; (2) poi distribution; (3) traffic condition. Afterward, we mapped the explicit feature vectors to the geographical spatial graph as the attributes of the corresponding node. Next, we utilized the graph embedding technique to fuse all explicit features and spatial relations in the context together to obtain the context embedding. Then we distinguished the excellent and terrible land-use configuration plans based on expert knowledge. Finally, the context embedding, excellent and terrible plans were input into our LUCGAN to learn the distribution of excellent plans. The LUCGAN can generate a suitable and excellent land-use configuration solution based on the context embedding when the model converges. Ultimately, we conduct extensive experiments to exhibit the effectiveness of our automatic planner.

## ACKNOWLEDGMENT

This research was partially supported by the National Science Foundation (NSF) via the grant numbers: 1755946, I2040950, 2006889.

## REFERENCES

[1] David Adams. 1994. *Urban planning and the development process.* Psychology Press.

[2] Samet Akcay, Amir Atapour-Abarghouei, and Toby P Breckon. 2018. Ganomaly: Semi-supervised anomaly detection via adversarial training. In *Asian conference on computer vision*. Springer, 622–637.

[3] Hugh Barton and Catherine Tsourou. 2013. *Healthy urban planning.* Routledge.

[4] Yunjey Choi, Minje Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, and Jaegul Choo. 2018. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 8789–8797.

[5] Ming Ding, Jie Tang, and Jie Zhang. 2018. Semi-supervised learning on graphs with generative adversarial nets. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. 913–922.

[6] J. Du, Y. Zhang, P. Wang, J. Leopold, and Y. Fu. 2019. Beyond Geo-First Law: Learning Spatial Representations via Integrated Autocorrelations and Complementarity. In *2019 IEEE International Conference on Data Mining (ICDM)*. 160–169.

[7] Yanjie Fu, Guannan Liu, Yong Ge, Pengyang Wang, Hengshu Zhu, Chunxiao Li, and Hui Xiong. 2018. Representing urban forms: A collective learning model with heterogeneous human mobility data. *IEEE transactions on knowledge and data engineering* 31, 3 (2018), 535–548.

[8] Yanjie Fu, Pengyang Wang, Jiadi Du, Le Wu, and Xiaolin Li. 2019. Efficient region embedding with multi-view spatial networks: A perspective of locality-constrained spatial autocorrelations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 906–913.

[9] Yaroslav Ganin, Tejas Kulkarni, Igor Babuschkin, SM Ali Eslami, and Oriol Vinyals. 2018. Synthesizing Programs for Images using Reinforced Adversarial Learning. In *International Conference on Machine Learning*. 1666–1675.

[10] Judy Hoffman, Eric Tzeng, Taesung Park, Jun-Yan Zhu, Phillip Isola, Kate Saenko, Alexei Efros, and Trevor Darrell. 2018. CyCADA: Cycle-Consistent Adversarial Domain Adaptation. In *International Conference on Machine Learning*. 1989–1998.

[11] Zhiyue Liu, Jiahai Wang, and Zhiwei Liang. 2020. CatGAN: Category-Aware Generative Adversarial Networks with Hierarchical Evolutionary Learning for Category Text Generation.. In *AAAI*. 8425–8432.

[12] Jari Niemelä. 1999. Ecology and urban planning. *Biodiversity & Conservation* 8, 1 (1999), 119–131.

[13] Samuel E Otto and Clarence W Rowley. 2019. Linearly recurrent autoencoder networks for learning dynamics. *SIAM Journal on Applied Dynamical Systems* 18, 1 (2019), 558–593.

[14] Yu-Ting Qiang, Yan-Wei Fu, Xiao Yu, Yan-Wen Guo, Zhi-Hua Zhou, and Leonid Sigal. 2019. Learning to generate posters of scientific papers by probabilistic graphical models. *Journal of Computer Science and Technology* 34, 1 (2019), 155–169.

[15] John Ratcliffe, Michael Stubbs, and Mark Shepherd. 2004. *Urban planning and real estate development.* Vol. 8. Taylor & Francis.

[16] Muhammad Sarmad, Hyunjoo Jenny Lee, and Young Min Kim. 2019. Rl-gan-net: A reinforcement learning agent controlled gan network for real-time point cloud shape completion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5898–5907.

[17] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. 2017. Adversarial discriminative domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 7167–7176.

[18] Pengyang Wang, Yanjie Fu, Jiawei Zhang, Xiaolin Li, and Dan Lin. 2018. Learning urban community structures: A collective embedding perspective with periodic spatial-temporal mobility graphs. *ACM Transactions on Intelligent Systems and Technology (TIST)* 9, 6 (2018), 1–28.

[19] Pengyang Wang, Xiaolin Li, Yu Zheng, Charu Aggarwal, and Yanjie Fu. 2019. Spatiotemporal Representation Learning for Driving Behavior Analysis: A Joint Perspective of Peer and Temporal Dependencies. *IEEE Transactions on Knowledge and Data Engineering* (2019).

[20] Pengyang Wang, Jiawei Zhang, Guannan Liu, Yanjie Fu, and Charu Aggarwal. 2018. Ensemble-spotting: Ranking urban vibrancy via poi embedding with multi-view spatial graphs. In *Proceedings of the 2018 SIAM International Conference on Data Mining*. SIAM, 351–359.

[21] Yunchao Zhang, Yanjie Fu, Pengyang Wang, Xiaolin Li, and Yu Zheng. 2019. Unifying inter-region autocorrelation and intra-region structures for spatial embedding via collective adversarial learning. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1700–1708.

[22] Yingxue Zhang, Yanhua Li, Xun Zhou, Xiangnan Kong, and Jun Luo. 2019. TrafficGAN: Off-Deployment Traffic Estimation with Traffic Generative Adversarial Networks. In *2019 IEEE International Conference on Data Mining (ICDM)*. IEEE, 1474–1479.

[23] Yingxue Zhang, Yanhua Li, Xun Zhou, Xiangnan Kong, and Jun Luo. 2020. CurbGAN: Conditional Urban Traffic Estimation through Spatio-Temporal Generative Adversarial Networks. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 842–852.

[24] Bo Zhu, Jeremiah Z Liu, Stephen F Cauley, Bruce R Rosen, and Matthew S Rosen. 2018. Image reconstruction by domain-transform manifold learning. *Nature* 555, 7697 (2018), 487–492.