Eurasia Journal of Science and Technology
ISSN: 2663-1016
DOI: https://doi.org/10.61784/ejst3081

AN ATTENTION-DRIVEN BUILDING CLUSTERING APPROACH TAKING SHAPE FEATURES INTO ACCOUNT

ZhiHua Liu
JiangXi University of Science and Technology, Ganzhou 341000, Jiangxi, China.
Corresponding Email: 6120220125@mail.jxust.edu.cn

Abstract: Spatial clustering is the basis of pattern recognition, which is of great significance to map generalization and map updating. Aiming at the problem that many clustering methods can not effectively use the building topology information and it is difficult to deal with high-dimensional data, an attention-driven fusion clustering method is proposed. The model consists of four modules : autoencoder (AE), graph convolutional neural network (GCN), attention-driven fusion module (AFGCN-H) and self-supervised module. AE and GCN are used to extract features from the original data describing the characteristics of buildings and the topological information of the spatial relationship of buildings, respectively. The AFGCN-H adaptively fuses the learning representations of different layers of the two modules. The self-supervised module will optimize the clustering label allocation through corresponding losses, adjust the network parameters, learn the features suitable for clustering, and improve the accuracy of clustering. This paper uses field data sets for clustering analysis, and compares the clustering effects of traditional k-means algorithm, DBSCAN algorithm and MST algorithm. The experimental results show that the proposed method is superior to the traditional spatial clustering algorithm in the results of vector building clustering analysis.
Keywords: Map generalization; Spatial clustering; Building group; Deep clustering

1 INTRODUCTION

Building complex synthesis is an important step in map generalization. Due to the complexity of the geometry and spatial distribution of buildings, it is a complex and difficult operation to synthesize buildings. As the first step of building generalization, spatial clustering can divide buildings into different types of groups according to the similarity and regularity of geometric, semantic and structural features of spatial entities without supervision, so that the objects of the same group are similar to each other. By constructing clustering tasks, it is possible to segment residential data with different characteristics and extract meaningful building spatial distribution patterns, aiming to improve complex building comprehensive tasks.

According to the clustering strategy, clustering algorithms can be divided into seven main types: partition-based algorithms, hierarchy-based algorithms, grid-based algorithms, graph-based algorithms, density-based algorithms, model-based algorithms and hybrid methods. The partition algorithm, such as k-means and k-medoid [1-2], divides the data into k categories, which requires prior expertise to determine k. However, the selection of k is vague in many cases, and such methods are difficult to find clusters of any shape, nor to determine outlier buildings. Hierarchical algorithms, such as CURE and BIRCH [3-4], can handle multi-scale spatial clustering problems, but they need to define termination conditions for them and they cannot find clusters of any shape. Many grid-based algorithms, such as fuzzy clustering and hierarchical clustering [5-6], show the problem of low efficiency and accuracy when the dimension of data is large. Graph-based algorithms, such as ASCDT, Minimum Spanning Tree (MST), and Delaunay triangulation [7-8], can identify buildings of any shape by taking into account the spatial structure and adjacency information of buildings. However, the gestalt principle, the calculation of the characteristic factors describing buildings, and the distribution of building data are not well understood.

Deep neural network is a nonlinear and deep-seated structural network with strong nonlinear transformation ability. It can extract the feature representation of these two types of data at the same time, and fuse the two types of features for more reliable feature extraction. This paper introduces the deep clustering method into the clustering of vector buildings, constructs a graph model for the building group, and abstracts a single building as a graph node. The deep neural network (DNN) and the graph convolutional neural network (GCN) are used to learn the node information and the graph structure information respectively, and the attention-driven method is used to dynamically fuse the features to construct the clustering model[9].

2 ATTENTION-DRIVEN FUSION CLUSTERING MODEL

The overall framework of the attention-driven fusion clustering model proposed in this paper is shown in Figure 1. The model mainly includes four parts: (1)Autoencoder (AE) learns the original geometric features of buildings; (2)The learning representation of geometric features is transferred to the graph convolutional neural network (GCN) to integrate geometric and structural information. (3) The attention mechanism is used to dynamically combine the features learned by each layer of autoencoder with the features learned by GCN; (4) The self-supervised module unifies the AE features and the fused features in a framework, and uses the relationship between them to optimize the clustering label allocation to learn the features suitable for clustering. In order to enable the model to process vector data, it is necessary to construct the adjacency matrix of the original data and calculate the feature factors used to abstract the building as a graph node[10].

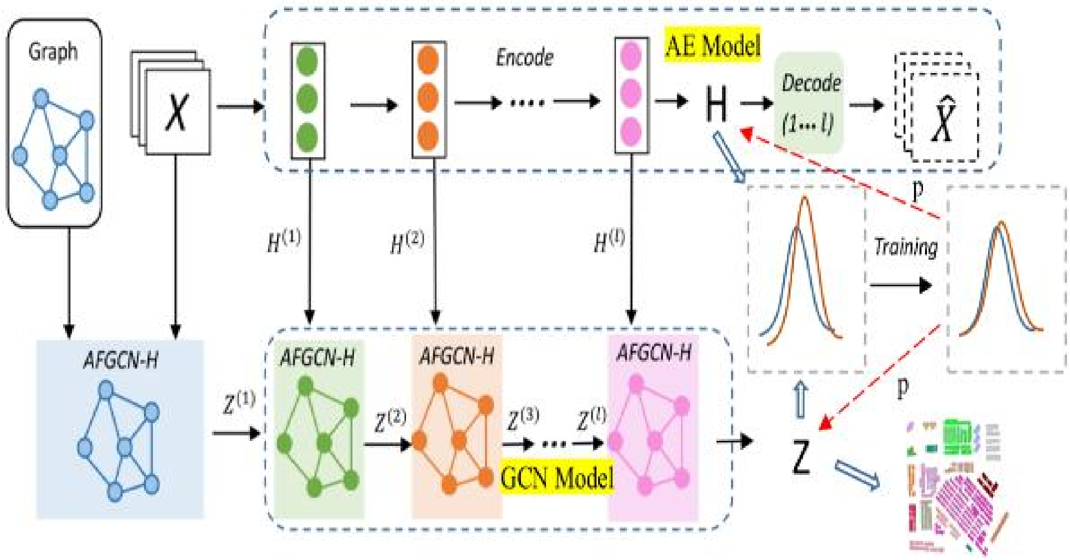


Figure 1 Building Attention-Driven Fusion Clustering Network

In Figure 1, X represents the characteristic matrix of the original input data, $\widehat{X}$ represents the reconstructed data, and Graph represents the adjacency matrix A, l represents the number of layers. $H^{(l)}$ represents the high-dimensional feature learning results of the AE autoencoder. The fusion $Z^{(l)}$ of GCN features and AE features is realized by the AFGCN heterogeneous fusion module (AFGCN-H) constructed by GCN. The network is self-trained by minimizing the KL divergence between the H distribution (orange curve) and the Z distribution (blue curve), and the update of the model is guided by the target distribution P.

2.1 Graph Structure Data

The construction of the graph model of the original building mainly includes obtaining the adjacency relationship and feature extraction between the buildings. The adjacency relationship can be obtained by constructing a constrained Delaunay triangulation. If there are triangles connected between buildings, the adjacency matrix value is set to 1, otherwise it is set to 0. In order to extract the features describing the nodes of the graph, this paper calculates 12 feature factors from the size, shape, direction and density of the building to describe the building. The data are organized into the original data $X \in R^{N\times d}$ and the adjacency matrix $A \in R^{N\times N}$, where N is the number of buildings and d is the number of node features[11].

2.2 Autoencoder Module

Since the composition of the original data is not complicated, this paper uses a basic autoencoder to learn the feature representation of the original data. Assuming that the autoencoder has a total L layer, the features learned at the l layer can be expressed as :

$$H^{(l)} = \emptyset\left(W_e^{(l)}H^{(l-1)} + b_e^l\right) \tag{1}$$

where $\emptyset$ is the activation function, $W_e^{(l)}$ and b_e^l are the weight matrix and bias term of the lth layer of the coding part, and $H^{(0)}$ is the original data X.

The encoded results are decoded by a fully connected network to reconstruct the original data, and finally the decoding result $\widehat{H}^{(l)}$ is calculated. The network structure of the encoder and decoder layers is completely symmetrical. Therefore, the final generated $\widehat{H}^{(l)}$ and $H^{(l)}$ have the same dimension, and their values should be as similar as possible. $\widehat{H}^{(l)}$ is calculated as follows :

$$\widehat{H}^{(l)} = \emptyset\left(W_d^{(l)}H^{(l-1)} + b_d^l\right) \tag{2}$$

$W_e^{(l)}$ and b_e^l represent the weight matrix and bias term of the layer l of the decoding part. The original data and reconstructed data can calculate the loss function :

$$\mathcal{L}_{res} = \frac{1}{2N}\sum_{i=1}^{N} \|x_i - \hat{x}_i\|_2^2 \tag{3}$$

2.3 Graph Convolutional Neural Network Module

The autoencoder can learn useful attribute feature information from the original data X, but this may ignore the topology information between building nodes.
The characteristic matrix X of the nodes in the graph and the adjacency matrix A representing the topological relationship between the nodes are used as the input of the GCN network. The GCN layer can be represented by the following formula :

$$Z^{l+1} = \emptyset\left(\widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}Z^{l}W^{l}\right) \tag{4}$$

Where $\widetilde{A} = A + I_N$ is the adjacency matrix of a self-connected undirected graph, I _ N is the identity matrix, $\widetilde{D}_{ii} = \sum_j \widetilde{A}_{ii}$ degree matrix, W^l is the trainable weighted matrix of each layer, $Z^l \in R^{N*D}$ represents the input of GCN at the l-layer, Z^0=X.
The GCN module is used to construct an automatic encoder to learn the spatial structure characteristics of the data in the encoding stage. The whole process can be summarized by the following formula :
$Z = \mathrm{GCN}(X, A)\#(5)$
$\mathrm{GCN}(\bullet)$ represents the graph convolution operation. In this module, the structural information Z^{l+1} learned by each layer of GCN can be obtained. By integrating the heterogeneous fusion module with the original node data reconstructed by AE, a more reliable and more distinguishable clustering data representation that can adapt to two different information can be obtained.

3 EXPERIMENTAL DATA AND PRETREATMENT

3.1 Experimental Region

In this paper, a relatively regular area of a street in Miami, the United States, is selected as the experimental area, with a total of 3234 buildings. This area contains clusters of different building modes, including linear mode, curve mode and grid mode. The distribution pattern of buildings is relatively rich. Through the experimental analysis of this area, the clustering performance of the model can be well tested, as shown in Figure 2.

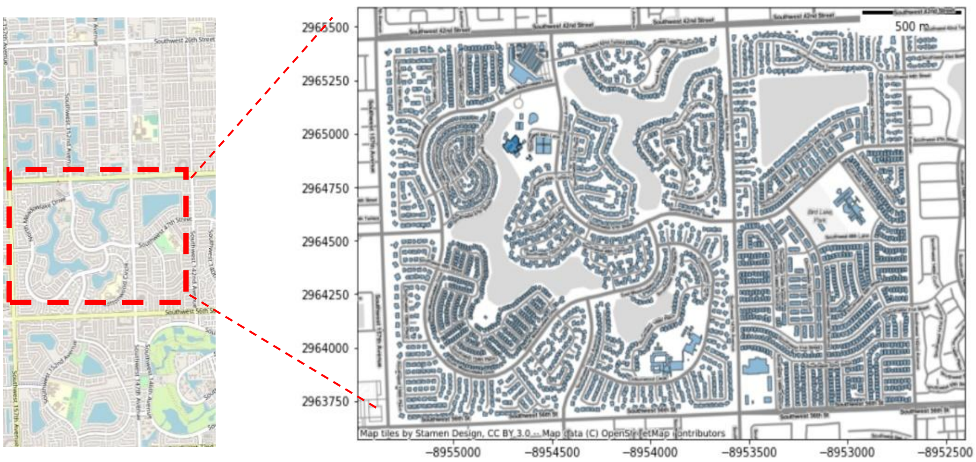
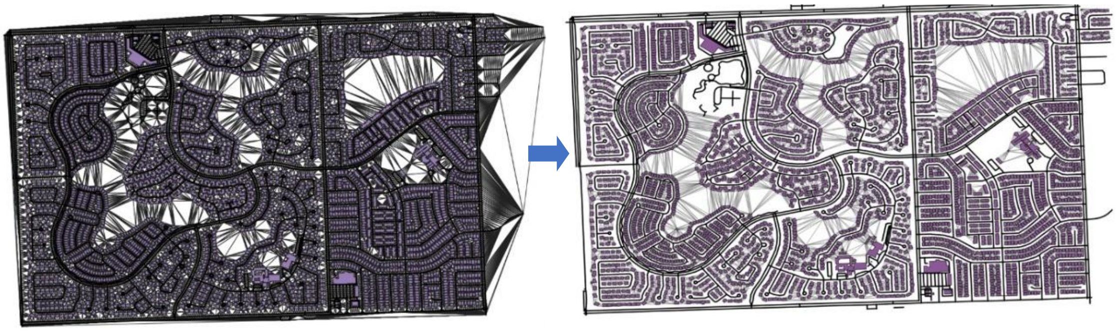


Figure 2 The Experimental Are

3.2 Construction of Geometric Graph Model of Building Group

Deep learning for vector building data should first build a geometric model of vector building data, and build a deep clustering model based on graph structure. An undirected graph can be defined as $G = (v, \varepsilon)$, where v and ε represent the adjacency relationship between the nodes of the graph and the two buildings, respectively. The construction of v will be discussed in the next section. In this paper, a constrained Delaunay triangulation (DT) subdivision is used to obtain its adjacency relationship and establish graph structure data for deep clustering. Firstly, the buildings are physically divided according to the road, and the buildings and roads are encrypted with identifiers at a distance of 10 meters, and the Delaunay triangulation is constructed according to the encrypted points, as shown in Figure 3a ; then delete the triangles related to the road encryption points according to the identifier, and remove the narrow and self-connected triangles to obtain an independent Delaunay triangulation of each block, as shown in Figure 3b; finally, according to the judgment of the first-order adjacent relationship of the buildings directly connected with the triangle, the adjacent matrix A of the building group is generated.
In the process of map generalization, the priority of roads is higher than that of general buildings. If two buildings are separated by roads in solid A, its value is 0. The adjacency matrix constructed in this way is more in line with human cognition and ensures efficiency.

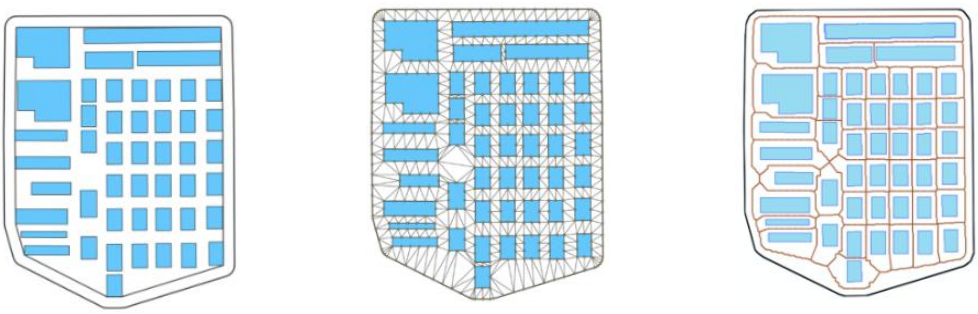


(a) (b)

Figure 3 Building Proximity: (a) Overall Delaunay triangulation; (b) Constrained integral Delaunay triangulation

3.3 Building Similarity Feature Factor

In order to construct the original data X, it is necessary to abstract the original building as a graph node. In order to better describe the geometric attribute information of the building, according to the principle of similarity and continuity, this paper constructs a quantitative model describing the building from the aspects of size, shape, direction and position of the building, so as to refine the grouping. The position is represented by the arithmetic mean of the building vertex. The size is described by the perimeter, area and the average distance from each vertex of the building to its centroid. According to the experimental results of Reference [12], the compactness, fractal dimension, concavity, perpendicularity and perimeter index are used to describe the shape of the building. The direction is represented by the minimum bounding rectangle (SBR) direction of the building polygon. Building density We divide the building group by DT, and generate the Voronoi like graph for the generated triangle according to the method of Reference [13].Each Voronoi like region in the graph is regarded as the affected area of the building, and its density is the ratio of the area of the building to the area of the affected area. In order to make the boundary buildings can also generate a completely surrounded Voronoi like region, we calculate its value according to the process of Figure 4, and the convex of the building group.



(a) (b) (c)

Figure 4 Calculation of Building Density Factor: (a) Convex hull buffer; (b) Construction constraint DT; (c) Generate a Voronoi-like diagram

The specific indicators are shown in Table 1. The node feature representation of the graph is calculated in the following way, that is, the feature matrix X. The subscript meanings involved in some of them are as follows : building (b), building convex hull (CH), and building minimum circumscribed matrix (SBR). In this paper, a total of 13 characteristic factors are calculated. In addition to the adjacent distance weighted to the adjacency matrix, the remaining factors are organized in the matrix form of $X \in R^{N\times d}$, N is the number of buildings, d = 12.

Table 1 Building Characteristic Parameters

parameter	Description index	formula	definition
position	centric position	$(x,y)=(\sum_{i=1}^{n}(x_i,y_i))/n$	The arithmetic mean of all vertex coordinates
size	area	A_b	area
	perimeter	P_b	The perimeter of each side and
	mean radius	$\overline{R}=\frac{1}{N}\sum_{i=1}^{n}R_i$	The average distance from the building vertex to the center position

shape	Compactness	$\mathrm{CI} = \frac{A_{pn}}{A_{EPC}} = \frac{4\pi A_b}{P_b^2}$	The area deviation between polygon and its isoperimetric circle
	fractal dimension	$\mathrm{FR} = 1 - \frac{\log A_b}{2 * \log P_b}$	Measuring edge roughness or smoothness
	concavity	$\mathrm{CNV} = \frac{A_b}{A_{CH}}$	The area deviation between the polygon and its convex hull is used to reveal the extent to which the polygon bends inward or outward.
	verticality	$\mathrm{REC} = \frac{A_b}{A_{SBR}}$	The area deviation between the polygon area and the SBR can reveal the degree of inward bending of the polygon.
	Perimeter index	$\mathrm{nPl} = \frac{P_{EAC}}{P_b} = \frac{2\sqrt{\pi A_b}}{P_b}$	Considering the compactness of polygon boundary
direction	SBR	$\alpha = \arctan\frac{x_i - y_i}{x_j - y_j}$	The longest edge direction of the minimum circumscribed rectangle
density	area ratio	$\mathrm{D} = \frac{S_B}{S_A}$	Deviation between building area and its affected area area
distance	visual distance	$D = \sum_{i=0}^{n} \frac{\lVert P_i P_{i+1} \rVert}{l} \lVert D_{i1} D_{i2} \rVert$	Triangle weighted distance, see [14]

4 EXPERIMENTS AND RESULT ANALYSIS

Through the above data preprocessing, the adjacency matrix A and the original data X can be obtained, and the number of categories K is determined by the number of blocks divided by the road. The original data is pre-trained by AE for 30 rounds to obtain the clustering centroid, and then the main model is used for 200 iterative training. The autoencoder size is set to : D-128-128-256-5, the graph convolution layer is set to : 128-128-256, and the average result of 10 experiments is selected as the final clustering result, as shown in Figure 5a.

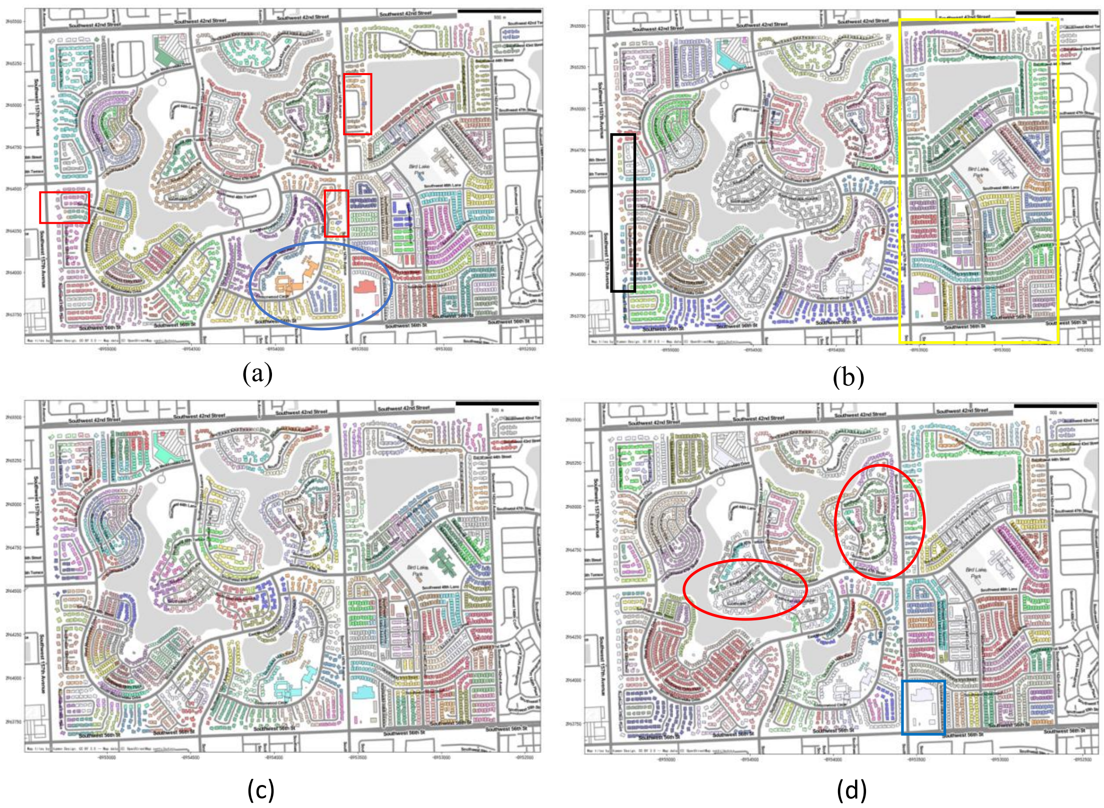


Figure 5 Clustering Results of Four Methods: (a) Proposed method; (b)mst; (c)k-means; (d)DBSCAN

There are 3234 buildings in the experimental area, and the results of the four clustering algorithms are shown in Figure 5. In Figure 5a, the accuracy of building division and human cognition can reach about 70 %. The whole experiment does not need too much parameter adjustment, and it is a complete end-to-end process from input to output. It can be seen that in the area formed by the road, the division effect of the building is ideal. For the linear pattern and the grid pattern distributed along the street, the deep clustering model can correctly identify (the lower right corner and the middle area), reflecting the importance of the topological information provided by the GCN module. If the established adjacency matrix is not segmented by the road, the clustering effect of the building groups arranged along the road and similar to the geometric properties of the building will be poor. In the scene of uniform and uneven building density, the clustering effect of the deep clustering model is also mild. This is due to the fusion of the feature representation of the two modules. The model can correctly classify the building groups according to the adjacency information and geometric information, as shown in the blue circle in Figure 5a.

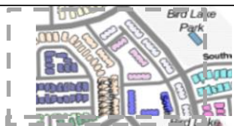
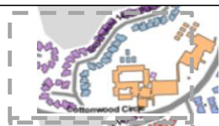
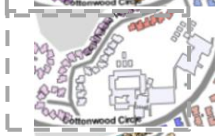
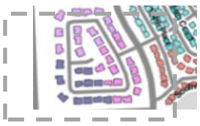
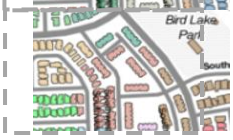
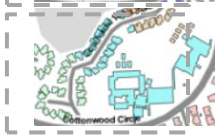
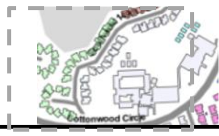
In addition, since the shape of the whole building does not change abruptly, one or two abrupt shapes in the cluster may be separated from the buildings that should belong to the same cluster. Therefore, the effective learning of the above building metrics is very important, as shown in the red rectangle. In general, the algorithm in this paper is valuable for building classification of large data sets, and the local effects are shown in Table 2.

Figure 5b describes the clustering effect of MST. It also uses DT to divide its adjacency relationship, and determines its optimal division threshold according to the dispersion. The specific implementation of the following three clustering algorithms is described in detail in Reference [14]. It can be seen that in the yellow rectangle, the division effect of MST is ideal, and buildings of any shape and density can be found. However, in many other areas, such as black rectangles, the recognition effect is very poor, and the simple linear pattern also has extreme errors. The reason is that the division threshold is a global parameter, which cannot adapt to each area of the whole, and the local area is more obvious. How to determine a reasonable ' pruning threshold ' has always been a problem faced by MST algorithm.

Figure 5c is the clustering effect of k-means. As a comparison, the k-means algorithm deals with the original building feature data, and cannot consider the adjacency topological relationship between buildings. The determination of its K value is the same as the method mentioned above. In the experimental process, k-means clustering has the highest efficiency, but it can be seen from the figure that its overall clustering effect is poor, especially for groups with linear patterns, such as buildings arranged along the road. In general, the shape of the cluster is similar to a sphere, and it is difficult to adapt to clusters of any shape and any density.

Figure 5d is the clustering effect of DBSCAN.Since the algorithm needs to determine the neighborhood radius eps and the minimum sample point min _ sample in the neighborhood radius, the optimal selection of the two is more difficult. The change of any party will lead to a sudden change in the number of clusters and the clustering effect. Therefore, this paper uses the cyclic iteration method to determine its optimal division parameters. The value of eps is 0.00031, and the value of min _ sample is 3, which can obtain an ideal effect diagram. In this experiment, the distribution of buildings is relatively uniform, and the density mutation is less. Therefore, the algorithm has achieved good results in the lower right corner of the experimental area, but in the middle area, as shown in the red circle, it should be continuous.

Table 2 Experimental Local Region Comparison

Clustering Algorithms	Region A	Region B	Region C
Methods			
K-means			
MST			
DBSCAN			

5 CONCLUSION

In this paper, deep learning is introduced to process vector data. The deep clustering model is built by using autoencoder and graph convolutional neural network. The attention mechanism is used to dynamically fuse the features learned by the two modules. The self-supervised module is constructed by student t distribution so that the whole clustering process is an end-to-end process. The experimental results show that after fully considering the structural and attribute characteristics of buildings, the deep clustering method can be better used for building clustering operations. Compared with traditional clustering algorithms, it has more advantages. It can not only deal with high-dimensional data, but also comprehensively consider the topological and geometric information of buildings.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Tardioli G, Kerrigan R, Oates M, et al. Identification of representative buildings and building groups in urban datasets using a novel pre-processing, classification, clustering and predictive modelling approach. Building and Environment, 2018, 140: 90–106.
[2] Vinod H D. Integer programming and the theory of grouping. Journal of the American Statistical Association, 1969, 64(326): 506–519.
[3] Guha S, Rastogi R, Shim K. CURE: An efficient clustering algorithm for large databases. ACM Sigmod Record, 1998, 27(2): 73–84.
[4] Zhu Y, Jiang Y. Optimization and parallelization of BIRCH clustering algorithm. Computer Engineering and Design, 2007, 28(18): 4345–4346, 4369.
[5] Li S, Zhang H, Wang L, et al. Meteorological drought risk zoning and analysis of winter wheat in Henan Province based on grid and fuzzy clustering. Journal of Henan Agricultural Sciences, 2020, 49(11): 172–180.
[6] Tang G, Xing C, Zhu L, et al. Rapid classification of moving surface fitting filter algorithm using hierarchical clustering. Engineering of Surveying and Mapping, 2021, 30(3): 32–40.
[7] Deng M, Liu Q, Cheng T, et al. An adaptive spatial clustering algorithm based on Delaunay triangulation. Computers, Environment and Urban Systems, 2011, 35(4): 320–332.
[8] Cai J. MST clustering of urban buildings considering Gestalt principles. Chang'an University, 2018. (Doctoral dissertation)
[9] Gong X, Fang Y. Comparative study on adaptability of settlement clustering analysis algorithms. Engineering of Surveying and Mapping, 2020, 29(5): 1–7.
[10] Pilehforooshha P, Karimi M. A local adaptive density-based algorithm for clustering polygonal buildings in urban block polygons. Geocarto International, 2020, 35(2): 141–167.
[11] Van der Maaten L, Hinton G. Visualizing data using t-SNE. Journal of Machine Learning Research, 2008, 9(Nov): 2579–2605.
[12] Basaraner M, Cetinkaya S. Performance of shape indices and classification schemes for characterising perceptual shape complexity of building footprints in GIS. International Journal of Geographical Information Science, 2017, 31(10): 1952–1977.
[13] Ai T, Zhang X, Zhou Q, et al. A vector field model to handle the displacement of multiple conflicts in building generalization. International Journal of Geographical Information Science, 2015, 29(8): 1310–1331.
[14] Meng N, Feng J, Jia Y. Comparative analysis of areal settlement clustering methods. Geomatics and Information Science of Wuhan University, 2022, 47(10): 1–5.