引文格式：徐蕾宇,孔博,肖天元,等. 基于图卷积神经网络的建筑物合并方法[J]. 地理空间信息, 2024, 22(6): 10-14. doi:10. 3969/j.issn. 1672-4623.2024.06.003

基于图卷积神经网络的建筑物合并方法

徐蕾宇[1]，孔　博[1]，肖天元[1]，余华飞[1]，艾廷华[1*]
（1. 武汉大学 资源与环境科学学院，湖北 武汉 430079）

摘　要：建筑物合并是地图综合中的复杂决策问题，基于规则的合并方法试图通过一系列认知参数和约束条件来解决“哪些合并”“如何合并”等问题，但是由于区域环境的复杂性、算子选择的多样性以及参量调整的多变性，基于规则的合并结果总是不尽人意。在人工智能技术支持下，将深度学习方法引入，则是从数据驱动视角通过样本训练以期突破规则方法的局限。本研究即是基于图卷积神经网络构建的一种建筑物分组及合并的方法。利用约束Delaunay三角网和最小生成树构建图结构，从图结构中提取建筑物节点特征和距离特征。通过图卷积层和点积运算得到的节点余弦相似度与距离特征拼接作为连接边特征，最后通过全连接层得到二分类结果。通过对比实验分析，本方法的ARI指数高于现有的方法（从0.674到0.800）。

关键词：地图综合；图卷积神经网络；建筑物合并
中图分类号：P283　　**文献标志码：**B　　**文章编号：**1672-4623（2024）06-0010-05

Building Aggregation Method Based on Graph Convolutional Neural Network

XU Leiyu[1], KONG Bo[1], XIAO Tianyuan[1], YU Huafei[1], AI Tinghua[1]
(1. School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China)

Abstract: Building aggregation is a complex decision-making problem in map generalization, where rule-based aggregation methods attempt to address the questions of “which to aggregate” and “how to aggregate” through a set of cognitive parameters and constraints. However, due to the complexity of regional environments, the diversity of operator selections and the variability of parameter adjustments, the results of rule-based aggregation are often unsatisfactory. With the support of artificial intelligence technology, introducing deep learning methods offers a way to overcome the limitations of rule-based approaches from a data-driven perspective. We presented a method for building grouping and aggregation based on graph convolutional neural network. We utilized constrained Delaunay triangulation and the minimum spanning tree to construct graph structure, extracting building node features and distance features from the graph. Then, we took the node cosine similarity obtained through graph convolution layers and dot product operations, concatenated with distance features as edge characteristics. Finally, we achieved a binary classification result through a fully connected layer. Through comparative experimental analysis, the ARI index of this method is higher than that of existing methods (from 0.674 to 0.800).

Key words: map generalization, graph convolutional neural network, building aggregation

地图综合是地图数据处理的经典问题，从传统的系列比例尺地图生成到现代智慧城市、数字孪生的多级瓦片地图生成，都需要地图综合技术的支持。地图综合算子是在特定条件下，通过选择、位移和简化等操作形式，完成地图图形的尺度变换。在长期的地图综合研究与实践中，产生了基于规则的综合方法。以建筑物多边形合并为例，基于规则的合并方法试图通过一系列认知参数和约束条件来解决“哪些合并”“如何合并”等问题。通常由于算子选择的多样性及参量调整的多变性，基于规则的合并结果总是不尽人意。随着人工智能技术的进步，特别是机器学习和深度学习的发展，为地图综合提供了新的解决方案[1]。这些技术能够从数据中学习和提取特征，为地图综合的自动化和智能化开辟了新的道路。在人工智能技术支持下，通过新的思路采用数据驱动的学习方法将是对规则方法的扩展。

在城市空间信息处理中，建筑物综合旨在从较大比例尺地图生成较小比例尺地图时解决空间冲突和拥挤，同时简化形态的细节表达[2]。建筑物合并作为地图综合的重要组成部分，其目的是将集群内的建筑物融合为单个对象以进行下一个更高级别的表示[3]。根据处理对象的不同，过去的建筑物合并研究可以分为

收稿日期：2024-02-27。
第一作者简介：徐蕾宇（2000—），硕士研究生，研究方向为地图综合，E-mail：xuleiyu@whu.edu.cn。
通信作者：艾廷华（1969—），博士，教授，研究方向为地图综合、空间分析，E-mail：tinghuaai@whu.edu.cn。

两类。第一种是针对栅格数据开发，包括使用形态学算子（扩展和侵蚀）的方法[4–6]，以及在2个垂直方向的扫描来填补建筑物之间的间隙实现合并的方法[7–8]。目前可以通过智能化的生成对抗网络对建筑物及其周围的栅格进行二分类实现建筑物化简、删除和合并等操作[9]。然而，这些基于图像的解决方法会在表示建筑物边缘时引入额外的噪声，或者在光栅化过程中，导致建筑物轮廓的位置精度损失[10–11]。第二种是针对矢量数据开发，具体可以分为四类，包括移位合并、泛洪合并、采样合并和连接对象合并[12]。本文采用最后一个合并策略即连接对象合并，通过图卷积神经网络对建筑物之间连接边进行分类，然后再使用ArcGIS进行后处理从而实现建筑物合并。

1　图结构下的特征分析

1.1　构建建筑物图结构

城市建筑物分布具有典型的群集特征，并表现出特定的模式形态，其分布格局体现了特定区域环境下的建筑风格，保持建筑物的分布模式是地图综合中的重要约束条件。面向建筑物模式的划分，应顾及道路通常作为自然边界的角色，将建筑物划分在地图空间内不同的区域，因此首先使用道路数据来划分建筑物。为了表达建筑群类不同目标间的邻域关系，需要通过特定几何构造模型探测识别建筑物间的“左邻右舍”关联，这里我们选用使用约束型Delaunay三角网（CDT）[13]来构建图结构，构建的图结构中，采用最小生成树的方法来简化图结构，从而加快模型的收敛速度。

1.2　特征提取

在执行建筑物合并操作之前，需要理解和量化建筑物的特征。Gestalt[14]心理学原则提供了一个理解建筑物形态的完形规则，其强调整体大于部分之和的观点，指出人们倾向于以有序、整体的方式感知视觉世界。例如，邻近性原则解释了为何紧密相邻的建筑物更容易被视为一组，而相似性原则帮助理解相似形状或尺寸的建筑物如何在视觉上形成聚集。这些原则促使制图员将相邻或相似的建筑物视为一个整体。建筑物合并的关键认知变量包括大小、方向、形状和密度等属性。此外，空间距离关系也是衡量地图上单个建筑物重要性的重要评估指标。在空间分析中，基本概念如空间自相关和区域间的空间异质性被反映在与建筑物分布相关的邻近度指标中。因此，为了将地理领域知识整合进建筑物合并的模型中，本研究定义了特定的特征指标，然后设计相应的几何计算方法来表示这些指标，如图1所示。其中，将单体建筑物特征构成的特征矩阵输入到图卷积神经网络中作为节点特征，而将连接边特征构成的特征矩阵作为距离特征。

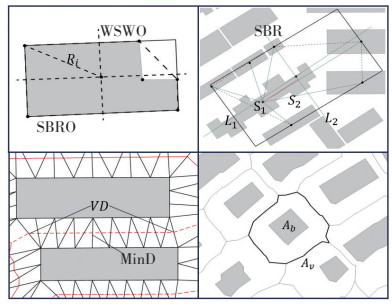


图1　节点特征和距离特征提取方法

2　GCN模型构建

2.1　图卷积神经网络

图卷积神经网络（GCN）[15]通过在图结构数据上应用卷积操作，为图中的节点学习有效的特征表示提供了一种高效的方法。在GCN中，每个节点的新特征表示是通过聚合其邻居节点的特征信息以及自身特征信息得到的。这一过程可以使用以下公式表达：

$$\boldsymbol{H}^{l+1}=\sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}\boldsymbol{H}^{l}W^{l}\right)$$

式中，$\boldsymbol{H}^{l}$ 为第 l 层的节点特征矩阵；$\boldsymbol{H}^{0}$ 为节点的初始特征矩阵；$\tilde{\boldsymbol{A}}=\boldsymbol{A}+I_N$ 为加上自环的邻接矩阵；$\boldsymbol{A}$ 为图的邻接矩阵；$\boldsymbol{I}_N$ 为单位矩阵；$\widetilde{\boldsymbol{D}}$ 为对角矩阵，其每个元素 $\tilde{\boldsymbol{D}}_{ii}$ 是节点 i 加上自环后的度数；$\boldsymbol{W}^{l}$ 为第 l 层的可学习权重矩阵；σ 为非线性激活函数，本文使用Relu作为激活函数。通过这种方式，GCN能够在每一层通过邻接节点的信息传播，学习到更高层次的节点特征表示。

2.2　建筑物合并模型

连接边的特征和连接边两端的建筑物节点特征相关，两端节点相似度越高，连接边越容易分类为合并。因此，需要计算连接边两端节点的余弦相似度：

$$H_{\text{sim}}^{uv}=\frac{H_{\text{node}}^{v}\cdot H_{\text{node}}^{u}}{\left\|H_{\text{node}}^{v}\right\|\cdot\left\|H_{\text{node}}^{u}\right\|}$$

$$H_e=\left[H_{\text{sim}}^{uv},H_{\text{dis}}^{uv}\right]$$

式中，H_{sim}^{uv} 为节点 u 和节点 v 之间的余弦相似度；H_{node}^{v} 为节点 v 的图卷积层输出；H_{node}^{u} 为节点 u 的图卷积层输出；$H_{\text{node}}^{v}\cdot H_{\text{node}}^{u}$ 为点积运算。将得到的节点间的余弦相似度 H_{sim}^{uv} 和距离特征 H_{dis}^{uv} 拼接作为连接边的特征 H_e，然后，通过一个全连接层来预测图中的

连接边，最终被预测为合并的边是输出层中概率高于未合并概率的连接边。将模型预测为合并的边以及由这些边连接的节点被识别为多个子图，每个子图都被分配一个唯一的类别编号。最后，使用ArcGIS中的“聚合面”工具，根据这些类别编号合并建筑物，实验流程如图2所示。

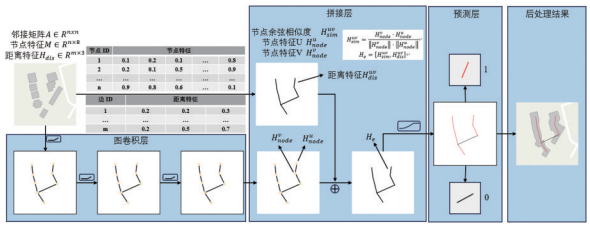


图2 连接边分类模型的详细流程

3 实验与结果

3.1 实验数据

为了评估模型在不同建筑群中合并的有效性，本实验使用了来自不同城市的建筑群数据。实验选择了广州和深圳2个地区。通过结合不同区域的建筑群，来全面评估模型对不同场景的适应性。数据集包括9 537栋建筑物和18 778条连接边，数据来源于百度地图。初始比例尺为1∶2 000，建筑物合并的目标比例尺为1∶5 000。关于数据集的划分，广州市番禺区的建筑物被指定为训练数据集和验证数据集。测试数据集则选取了深圳市龙华区2个不同区域的建筑群。广州的训练数据集和验证数据集包含各种不同的建筑物排列模式，深圳的测试数据集1与测试数据集2相比，包含更多的建筑物和连接边。表1列出了每个数据集的建筑物和标注连接边的具体数量。

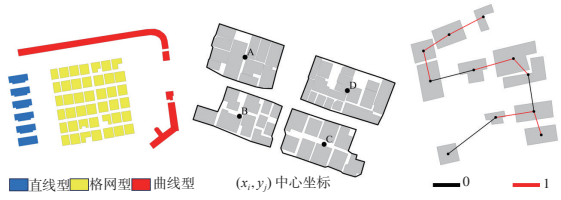
表1 数据集简介

数据集	城市	建筑物数量	连接边数量	街区数量
番禺区训练集	广州	2 234	4 588	41
番禺区验证集	广州	946	1 650	19
龙华区测试集1	深圳	3 811	7 504	52
龙华区测试集2	深圳	2 546	5 036	23

本实验中，使用2个卷积层，学习率、正则化权重、训练次数分别为0.01，5e−4和200，并选用Adam作为优化器。

3.2 图标签制作

在现有的建筑物综合研究中，节点通常被指定为建筑物，这些建筑物呈现出不同的形态，如直线型、格网型和曲线型[16]，如图3a所示。根据这些模式，单个建筑物被赋予不同的标签。另一种节点标注方法是将每个组内的建筑物与组轮廓的中心点联系起来，从而有效地为每个建筑物分配一个二维向量[17]，如图3b所示。在本研究中，我们的方法进行了创新，标注的是图中建筑物节点之间的连接边。具体来说，同一建筑群内的连接边标记为1，而连接不同建筑群的连接边标记为0，如图3c所示。



a 建筑群模式 b 建筑群中心坐标 c 建筑群连接边

图3 建筑组群标签

3.3 评价指标

为了评估建筑群合并方法的质量，选择合并后建筑群和初始建筑群之间调整兰德指数（ARI）作为评判指标。建筑标签和预测结果将建筑物分类为2组：$X=(x_1,x_2,\cdots,x_l)$ 作为标签类别，和 $Y=(y_1,y_2,\cdots,y_p)$ 作为预测类别。评估的ARI指数公式如下：

$$\mathrm{ARI}=\frac{\sum_{ij}\binom{n_{ij}}{2}-\frac{\left[\sum_i\binom{n_i}{2}\sum_j\binom{n_j}{2}\right]}{\binom{n}{2}}}{\frac{1}{2}\left[\sum_i\binom{n_i}{2}+\sum_j\binom{n_j}{2}\right]-\frac{\left[\sum_i\binom{n_i}{2}\sum_j\binom{n_j}{2}\right]}{\binom{n}{2}}}$$

式中，n_{ij} 为 x_i 与 y_j 之间交集集合中的建筑物数量；n_i 和 n_j 分别为 x_i 和 y_j 中的建筑物数量；$\binom{m}{n}$ 为 $\frac{n!}{(n-m)!\,m!}$。

3.4 建筑物合并结果

图4展示了使用番禺区数据进行连接边分类的训练过程。值得注意的是，模型在大约100次迭代后趋于稳定，本方法在训练和验证数据集上的准确率都超过了80%。对于每个数据集，分类模型都只被用来预测连接边是否应该合并，然后使用ArcGIS来聚合预测为合并边所连接的建筑物。所使用的连接边分类模型展示了其在合并不同区域的建筑群方面的有效性。

图5显示了分类模型在深圳第一个测试数据集上取得的建筑物合并结果。分类模型有效地学习了建筑物合并的基本规则，并成功合并了大部分建筑物。图6展示了深圳第二个测试数据集的合并结果。结果表明，所提出的连接边分类模型在不同区域都表现出有效的泛化能力。这种聚合方法并非基于预定义的

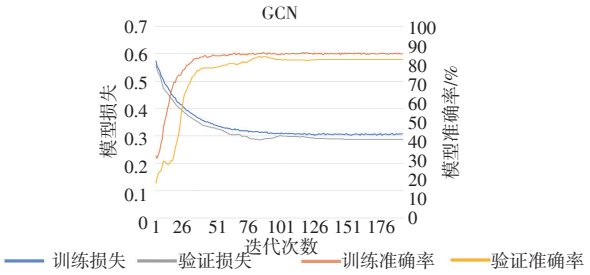


图4　模型的训练损失、验证损失、训练准确率和验证准确率随时间的变化

建筑物模式，而是依靠模型自主学习建筑物合并通用规则。

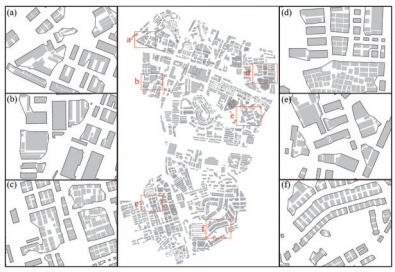


图5　模型在测试集1上的合并结果



图6　模型在测试集2上的合并结果

3.5　对比实验

在对比实验中，我们评估了2种简单易行的合并方法：DBSCAN和基于ArcGIS的聚合面方法（简称为ArcGIS方法）。针对这2种方法涉及到的参数，本研究参照现有研究[18-19]，并经过多次实验，确定了适宜的参数：DBSCAN方法的全局邻域搜索半径（eps）为23 m，最小点数（minPts）为2，ArcGIS方法的搜索半径为5 m。另外，为了和同种机器学习算法进行比较，在模型中用两层多层感知机（MLP）层替代了原有的两层图卷积层，而保持其他部分不变。表2汇总了使用这4种方法在所有街区数据集上的平均调整兰德指数（ARI）结果。基于图卷积神经网络的方法在2个测试集上均实现了最高的ARI值，分别为0.800和0.674，显示出最优性能。基于ArcGIS的方法，在第一个数据集上的ARI值为0.731，在第二个数据集上为0.600。该方法性能的降低可能源于为不同地区选择适当的聚合距离参数的挑战性，需要根据每个地区的建筑物调整参数。此外，依赖距离进行聚类可能不适用于所有场景。相比之下，DBSCAN方法在这些数据集上的ARI表现未达到预期，原因是该方法将建筑物聚类处理成点聚类，未考虑建筑物的空间形状和方位。将图卷积层更换为多层感知机后，由于模型不再处理图结构信息，从而未能充分考虑邻居节点的特征，这导致在第一个测试集上ARI值下降至0.577，在第二个测试集ARI值降至0.539。

表2　在2个数据集上使用不同方法获得的ARI指数

方法	GCN	ArcGIS	DBSCAN	MLP
测试集1的平均ARI	0.800	0.731	0.601	0.577
测试集2的平均ARI	0.674	0.600	0.526	0.539

图7展示了在2个数据集上使用不同方法获得的ARI指数分布。对于第一个测试数据集，4种方法的ARI指数分布存在显著差异，许多区域的指数集中在0.1以下和0.9以上。值得注意的是，使用MLP来代替图卷积层的方法得分低于0.1的区域比例最高，占总区域的40%。在第二个测试数据集上，基于GCN的方法表现突出，73%的区域分组性能高于0.5。相比之下，另外3种方法分别只有65%、39%和43%的区域得分超过0.5。这些对比实验结果凸显了所提出的方法在不同区域的建筑聚合方面的卓越性能。

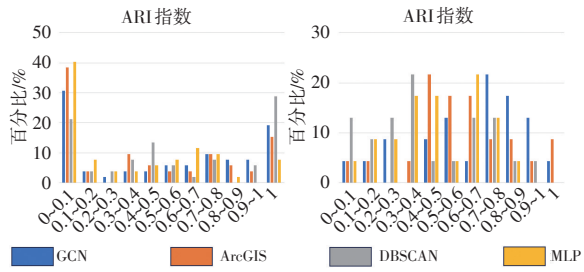


图7　在2个数据集上使用不同方法获得的ARI指数百分比

4　结　语

建筑物合并是地图综合中的一个关键任务。本文方法利用图卷积神经网络（GCN）来构建一个边分类框架，旨在预测建筑物之间的连接边从而实现建筑物合并。实验结果表明，本文提出的方法充分考虑了单个建筑的内在特性和其周围空间分布特征，建筑物合并效果相对传统方法更好。此外，本文还评估了边分类模型在不同区域的泛化能力。实验结果证明了模型可以实现在不同地区的建筑物合并，泛化性很好。

参考文献

[1]　艾廷华. 深度学习赋能地图制图的若干思考[J]. 测绘学报，2021,50(9):1 170-1 182

[2]　Yan H, Weibel R, Yang B. A multi-arameter Approach to Au-

tomated Building Grouping and Generalization[J]. Geoinformatica,2008,12:73-89

[3] Ware J M, Jones C B, Bundy G L. A Triangulated Spatial Model for Cartographic Generalisation of Areal Objects[C]//Spatial Information Theory A Theoretical Basis for GIS: International Conference COSIT'95 Semmering, Austria, September 21-23, 1995 Proceedings 2. Springer Berlin Heidelberg,1995: 173-192

[4] Schylberg L. Cartographic Amalgamation of Area Objects[J]. International Archives of Photogrammetry and Remote Sensing,1993,29: 135-135

[5] Su B, Li Z, Lodwick G, et al. Algebraic Models for the Aggregation of Area Features Based Upon Morphological Operators[J]. International Journal of Geographical Information Science,1997,11(3): 233-246

[6] Cámara M, López F. Mathematical Morphology Applied to Raster Generalization of Urban City Block Maps[J]. Cartographica: The International Journal for Geographic Information and Geovisualization,2000,37(1): 33-48

[7] Ai T, Zhang X. The Aggregation of Urban Building Clusters Based on the Skeleton Partitioning of Gap Space[J]. The European Information Society: Leading the Way With Geo-Information,2007: 153-170.

[8] Cheng B, Liu Q, Li X. Local Perception-based Intelligent Building Outline Aggregation Approach with Back Propagation Neural Network[J]. Neural Processing Letters, 2015,41(2): 273-292

[9] Feng Y, Thiemann F, Sester M. Learning Cartographic Building Generalization with Deep Convolutional Neural networks[J]. ISPRS International Journal of Geo-information, 2019,8(6): 258

[10] Zhou Z, Fu C, Weibel R. Move and Remove: Multi-task learning for Building Simplification in Vector Maps with a Graph Convolutional Neural Network[J]. ISPRS Journal of Photogrammetry and Remote Sensing,2023,202: 205-218

[11] He X, Zhang X, Yang J. Progressive Amalgamation of Building Clusters for Map Generalization Based on Scaling Subgroups[J]. ISPRS International Journal of Geo-information,2018,7(3): 116

[12] Regnauld N. Algorithms for the Amalgamation of Topographic Data[C]//Proceedings of the 21st International Cartographic Conference, Durban,South Africa. 2003,1016
graph convolutional networks[J]. arXiv preprint arXiv: 1609.02907,2016

[13] Shewchuk J R. Triangle:Engineering a 2D Quality Mesh Generator and Delaunay Triangulator[C]//Workshop on Applied Computational Geometry. Berlin,Heidelberg: Springer Berlin Heidelberg,1996:203-222

[14] Peura M, Iivarinen J. Efficiency of Simple Shape Descriptors[C]//Proceedings of the Third International Workshop on Visual form 1997,5:443-451

[15] Xiang Z,Tinghua A,Stoter J. The Evaluation of Spatial Distribution Density in Map Generalization[J]. The International Archives of the Photogrammetry, Remote Sensing and SpatialInformation Sciences,2008(6) 37-43

[16] Kipf T N, Welling M. Semi-supervised Classification with Graph Convolutional Networks[J]. arXiv Preprint arXiv: 1609. 02907,2016

[17] Zhao R,Ai T,Yu W,et al. Recognition of Building Group Patterns Using Graph Convolutional Network[J]. Cartography and Geographic Information Science,2020,47(5):400-417

[18] Yan X, Ai T, Yang M, et al. A Graph Deep Learning Approach for Urban Building Grouping[J]. Geocarto International,2022,37(10):2 944-2 966

[19] 钱海忠,武芳,谭笑,等. 基于ABTM的城市建筑物合并算法[J]. 中国图象图形学报,2005,10(10):1 224-1 233