引文格式：罗浩,廖顺华,周何帆,等. 城市全域道路多级自动选取方法[J]. 地理空间信息, 2025, 23(6):121-125. doi:10. 3969/j.issn. 1672-4623.2025.06.027

城市全域道路多级自动选取方法

罗　浩[1,3]，廖顺华[2*]，周何帆[1,3]，程灵雅[1,3]，邓维熙[1,3]，张　璇[1,3]
（1. 自然资源部第三地理信息制图院，四川 成都 610100；2. 广西壮族自治区地图院，广西 南宁 530200；
3. 自然资源部数字制图与国土信息应用重点实验室，四川 成都 610100）

摘　要：道路多级选取是地图综合的核心问题。针对现有方法多局限于小范围区域的问题，创新性地提出融合自适应特征与多级分类一体化的城市全域道路选取方法。通过多属性约束将路网划分为stroke单元，构建自适应空间特征和语义特征，结合GraphSAGE深度学习模型实现多级分类训练。实验表明，该方法在城市全域道路选取中实现了较高选取精度，并具有良好的跨区域泛用性。

关键词：道路网；道路选取；GraphSAGE；多级分类一体化
中图分类号：P283　　文献标志码：B　　文章编号：1672-4623（2025）06-0121-05

Multi-level Automatic Selection Method for Urban Roads

LUO Hao[1,3], LIAO Shunhua[2], ZHOU Hefan[1,3], CHENG Lingya[1,3], DEN Weixi[1,3], ZHANG Xuan[1,3]
(1. The Third Geoinformation Mapping Institute of Ministry of Natural Resource, Chengdu 610100, China;
2. Guangxi Zhuang Autonomous Region Institute of Cartography, Nanning 530200, China;
3. Key Laboratory of Digital Mapping and Land Information Application, Ministry of Natural Resources, Chengdu 610100, China)

Abstract: Multi-level road selection is a core issue in map generalization. To address the limitations of existing methods that are often confined to small-scale areas, we innovatively proposed an urban road selection method integrating adaptive features and multi-level classification. We segmented the road network into stroke units through multi-attribute constraints, constructed adaptive spatial and semantic features, and leveraged the GraphSAGE deep learning model to achieve multi-level classification training. Experimental result demonstrates that the proposed method achieves high selection accuracy in urban road selection and has strong cross-region generalizability.

Key words: road network, road selection, GraphSAGE, multi-level classification integration

道路选取作为道路网综合的核心问题，由于涉及对众多道路目标之间全局关联的分析与判断，其复杂程度相对更高。道路选取的合理性直接关乎地图综合的精度与质量[1]。因此，深入探究道路网选取方法，不仅具有重要的理论价值，而且在实际应用中拥有广阔的发展前景。

1　道路自动选取方法

1.1　传统道路自动选取

传统道路选取方法是道路网自动选取研究的早期探索方向，其核心思想是基于道路的几何特性、语义属性以及拓扑关系对道路进行筛选与综合。早期研究多集中于基于道路等级的简单选取，主要考虑道路的长度、等级等特征，将每条道路作为单独的个体进行分析和处理。这种方法对制图员的经验依赖较强，选取结果具有一定的主观性，难以适应复杂的道路网络场景[2-4]。

1.2　基于深度学习的道路自动选取

随着深度学习技术的飞速发展，其逐渐成为遥感和地图制图等领域的重要技术之一[5-7]。相比于传统机器学习，深度学习通过多层神经网络结构实现特征的自动提取，具备更强的特征挖掘和表达能力，同时免去了复杂的人工特征设计。因此，这一技术在地图综合中得到了成功应用[8-9]。

尽管结合深度学习的自动选取方法相比传统自动选取方法，在城市全域大规模道路网任务中，选取效果已经有了显著提升，然而与制图员选取的结果相比，在路网密度、连通度、美观性方面都有差距。

为了应对上述问题，本文创新性地提出一种融合多级分类一体化与自适应特征，并结合图采样聚合神经网络的道路自动选取方法。

第一作者简介：罗浩（1992—），硕士，高级工程师，主要从事地图制图与地理信息系统研发工作，E-mail：932238064@qq.com。
通信作者：廖顺华（1974—），本科，高级工程师，主要从事地图制图与地理信息系统应用工作，E-mail：158747079@qq.com。

2 研究方法

本文研究内容主要分为4个部分：道路stroke构建，模型框架，多级分类一体化，模型训练与测试。整体技术流程如图1所示。

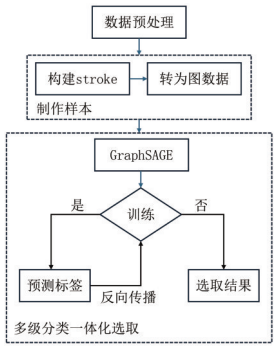


图1 技术路线流程图

2.1 道路stroke构建

进行道路stroke构建，首先要进行道路数据提取和预处理，预处理包括拓扑检查、悬挂伪节处理、实部相交断线等。

接下来构建stroke模型如图2，其中最后构建出的道路stroke为示意图(与a、b、c的几何意义不同)，不同颜色代表stroke满足不同约束条件。

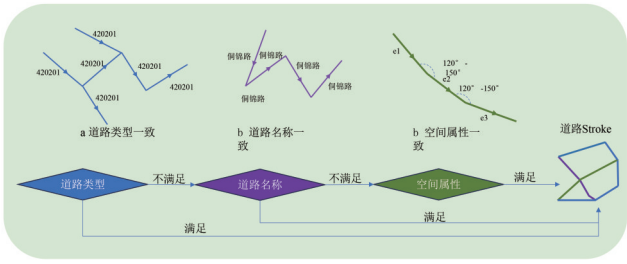


图2 道路Stroke构建过程

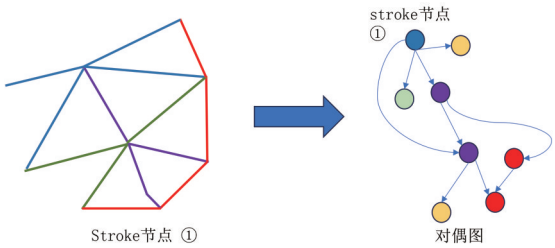


图3 stroke转化为图结构

道路stroke构建完毕后，需将所有道路stroke转化为对偶图结构，才能用于后续图神经网络模型的训练，过程如图3所示。对偶图是一种数据结构，包含图节点和边2个基本要素，其中道路stroke作为图节点，道路stroke之间的关联信息，作为边。这个过程可以作如下描述：定义 $G=\{V,E,\boldsymbol{W}\}$ 为一个图，其中 $V=\{v_1, v_2, v_3,\cdots, v_n\}$ 代表图中 n 个节点，E 为节点之间连接的边的集合，$\boldsymbol{W}$ 为 $n\times n$ 的邻接矩阵，存储每一对节点之间的关联信息权重，若节点之间存在关联信息，则为1，代表有关联；反之，则为0，代表无关联，在本研究中采用欧氏距离作为关联信息，通过计算2个stroke几何中心之间的欧氏距离，与设定好的阈值进行比较，大于阈值的判定为距离过远，代表无关联，小于阈值的判定为有关联，几何中心计算公式如下：

$$C_j=\left(\frac{1}{n_j}\sum_{i=1}^{n_j}x_i,\frac{1}{n_j}\sum_{i=1}^{n_j}y_i\right) \tag{1}$$

式中，C_j 为第 j 个stroke的几何中心。欧氏距离的计算公式如下：

$$d_{jk}=\sqrt{\left(x_c^j-x_c^k\right)^2+\left(y_c^j-y_c^k\right)^2} \tag{2}$$

式中，2个stroke的中心坐标分别为 (x_c^j,y_c^j) 和 (x_c^k,y_c^k)；d_{jk} 为第 j 个stroke和第 k 个stroke之间的欧氏距离。

本研究在传统空间特征的基础上，引入了具有自适应性的语义特征和空间特征，如图4所示，以成都市为例，选中最中心、最左下方和最左上方的A，B，C 3个道路stroke，其中 d_{AB}，d_{AC} 和 d_{BC} 分别表示3点之间的距离，从图上可以看出3个stroke之间的距离就是自身与其他stroke之间的最大欧氏距离，且处于边缘区域的A和C之间欧氏距离明显大于B点与两点之间。此外，除了空间分布，长度以及道路类型也有区别，为了量化这种差异，提升方法面对城市全域道路选取的效果和跨城市的泛用性，本文提出了相对位置因子。首先按照公式（1）计算道路stroke的几何中心 C_j，根据公式（2）依次计算每个道路stroke与其他所有道路stroke的欧氏距离，然后算出每个道路stroke与其他所有道路stroke的距离极差：

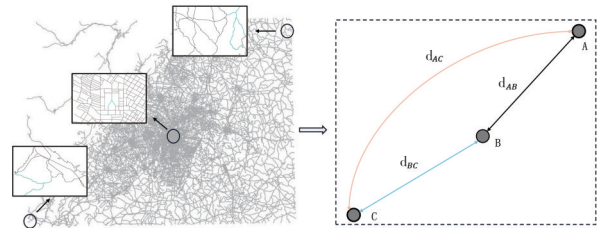


图4 不同区域道路stroke空间分布示意图

$$\Delta d_j=\mathrm{Max}\left(d_j\right)-\mathrm{Min}\left(d_j\right) \tag{3}$$

再以道路stroke长度和道路类型等级为综合权重，分别表示为 w_l 和 w_g，计算得到相对位置因子 δ：

$$\delta=\log w_l\times\log w_g{}^2\times\Delta d_j\times10^{-4} \tag{4}$$

最终，本文选取的特征因子有长度、直曲比、连通度、道路类型等级、相对位置因子，表1中展示了

2 个 stroke 样本特征示例。

表 1　stroke 特征示例

特征	长度	直曲比	连通度	道路类型等级	相对位置因子
stroke1	2 724.65	0.847	6	13	4.32
stroke2	504.67	0.99	3	14	3.27

2.2　模型框架

针对图数据的分类训练，图神经网络（graph neural networks，GNN）[10]有着天然的优势，可以充分挖掘不同节点之间的潜在关系，用于下游各种任务。在图神经网络的基础上，衍生出的图采样聚合网络（graph sample and aggregation，GraphSAGE）[10-12]，在各种基于图数据的分析任务中，表现出了更加优异的性能，同时非常适合大规模图数据的节点表示学习，因此，本研究中采用了这个模型，GraphSAGE 的模型结构如图 5 所示。

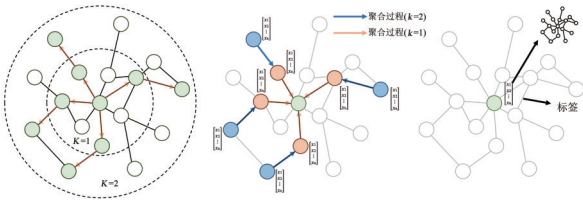


图 5　图采样聚合网络结构图

模型运算流程包含特征聚合和线性计算 2 个部分，特征聚合会随机选择目标节点的 K 个邻居节点，采用平均聚合函数，完成信息聚合，再通过线性层，将聚合后的特征映射到用于分类的特征空间中，完成输出。具体流程可以如下表示，首先给定图数据：

$$G=\{V,E\} \tag{5}$$

式中，V 为节点集合；E 为边集合；每个节点 $v\in V$ 有特征向量 $\boldsymbol{h}_v^{(0)}$，特征聚合阶段，先进行邻居节点的采样，获得采样子集：

$$N_v\subseteq\{u|(v,u)\in E\} \tag{6}$$

式中，N_v 为节点 v 的采样邻居节点，再进行特征信息的聚合与节点特征更新，其具体计算过程如公式（7）与公式（8）所示：

$$\boldsymbol{h}_v^{(k)}=\mathrm{MEAN}^{(k)}\left(\left\{\boldsymbol{h}_v^{(k-1)}|u\in N_u\right\}\right) \tag{7}$$

$$\boldsymbol{h}_v^{(k)}=\sigma\left(\boldsymbol{W}^{(k)}\cdot\mathrm{CONCAT}\left(\boldsymbol{h}_v^{(k-1)},\boldsymbol{h}_{N_v}^{(k)}\right)\right) \tag{8}$$

式中，$\boldsymbol{W}^{(k)}$ 为第 k 层的可学习权重矩阵；$\sigma(\cdot)$ 为非线性激活函数，经过 K 次传播后，得到节点的最终特征表示 $\boldsymbol{h}_v^{(K)}$，最后通过全连接层的线性计算得到最终预测输出，具体公式（9）如下：

$$\mathrm{pre}=FC\left(W\times\boldsymbol{h}_v^{(K)}+b\right) \tag{9}$$

2.3　多级分类一体化

本研究提出了多级分类一体化的选取策略，提升方法在城市全域大规模道路网多种比例尺连续性选取中的精度。每种标准下的道路选取，是个简单的二分类任务。但不同比例尺下，道路选取的标准并不相同，此前基于深度学习的自动选取方法中，每种标准下的选取相对独立，意味着每一次都要独立训练，影响效率。同时，由于独立训练，模型每一次面对的样本特征空间都是陌生的，很难进一步提升选取精度。

在道路选取中，为保证跨比例尺要素的选取一致性，小比例尺的道路网通常是从大比例尺道路网中逐级选取出来的。因此，可以将多种比例尺的选取融合到一起，以提升效率，同时采用连续训练的方式，每一次训练都是基于上一次训练的模型参数，从而让模型不断深入拟合样本的特征空间，提升整体精度。在选取时按照由大比例尺到小比例尺的顺序进行[10]，整体策略如图 6 所示。

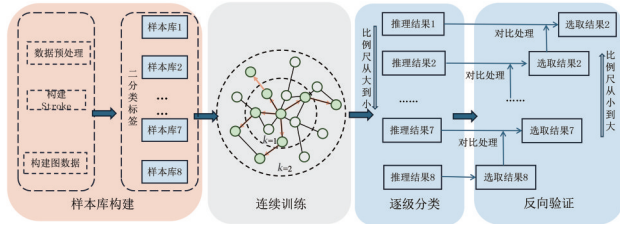


图 6　多级分类一体化策略

2.4　模型训练策略

本实验基于 NVIDIA Geforce RTX 3090 GPU 的实验平台。对于道路 stroke，计算了长度等 3 个空间特征以及道路类型等级和道路相对位置因子 2 个自适应特征作为图节点特征。在进行模型训练之前，对所有节点特征进行了归一化。每个城市数据集都采用五折交叉验证的方式，对样本进行划分。

所用训练模型为 GraphSAGE，聚合函数采用平均聚合，并设置随机采样的邻居节点数量为 8。模型学习率设置为 0.002，采用 Adam 优化器，设置最大训练次数为 20 000，并采用 StepLR 作为学习率调度器，控制学习率每 30 次便衰减，衰减率为 0.95，损失函数采用交叉熵损失函数[13]。根据采用了早停机制控制训练过程，忍耐度设置为 300，使用损失函数作为监控指标，即模型在 300 次训练后，损失函数不继续下降，则跳出训练，保存模型。

实验数据中某些比例尺下，例如 1∶1 000 000 比例尺中，选取的样本比例远远低于不选取样本，为了

缓解模型在训练过程中，样本不平衡带来的预测偏向，实验中通过重新设定正类的判定阈值（初始默认为0.5），即判定为选取的阈值来实现。通过大量实验，最终确定比例尺由大到小的阈值设定分别为[0.05, 0.25, 0.3, 0.2, 0.5, 0.6, 0.8]。

3 实验结果

本研究选取四川省成都市、广西壮族自治区南宁市和三江侗族自治县的1∶5 000道路数据用作实验，数量分别为成都市41 547条（去除伪结点前90 903条），南宁市46 654条（去除伪结点前54 654条），三江侗族自治县7 995条（去除伪节点前13 767条）。最终成都市样本库中图节点数量，即道路stroke数量为14 987个，南宁市样本库中图节点数量为16 786个，三江侗族自治县样本库中图节点数量为2 677个，每个城市的数据只用于自身的实验。

实验中，采用精确率（Precision）、召回率（Recall）、F1-score和选取一致率作为模型的评估标准，前3个指标在训练中以stroke为基础直接计算，可以帮助分析模型与特征存在的缺陷和不足，计算公式如表2所示，选取一致率以道路本身为基础计算，可以十分直观地呈现选取效果。每个比例尺结果数据代表该比例尺中二分类选取的实验结果。

表2 指标计算公式

指标	公式
Precision	$P=TP/(TP+FP)$
Recall	$R=TP/(TP+FN)$
F1-score	$F1=(2\times P\times R)/(P+R)$
选取一致率	（正确选取/应选取+正确不选取/不应选取）/2

3.1 整体选取效果

本实验完成了四川省成都市、广西壮族自治区南宁市和三江侗族自治县3个城市的多级选取实验。为了同时展现在小比例尺和大比例尺下的选取效果，一共选取了1∶10 000、1∶500 000和1∶100 0000 3个比例尺。其中成都市比例尺1∶10 000的被选取，没被选取的道路数量的占比，小于南宁市和三江侗族自治县。

3个城市的8个级别比例尺的详细选取结果，效果最好的是1∶100 000到1∶100 0000到比例尺下的选取，选取一致率能达到92%以上，效果最差的是1∶25 000比例尺下的选取，选取一致率只有80%左右。从结果中可以看到：选取一致率呈现先降后升的趋势，四项指标从1∶5 000到1∶25 000逐步降低，从1∶25 000到1∶1000 000逐步升高。这是由于从1∶5000到1∶25 000比例尺过渡中，道路网的异构性显著增强，次要道路与主干道的几何相似性导致特征空间重叠度增加，模型难以建立清晰的决策边界。而在1∶25 000比例尺之后，道路要素的语义特征主导性显著增强，保留的道路均为高等级主干道，其几何特征和拓扑特征呈现显著可分性，使得选取精度有所提升。

3.2 对比常规深度学习方法的结果

本实验还选取此前常规深度学习方法作为比对方法，常规深度学习方法，提取3种传统空间特征，即stroke长度、stroke直曲比、连通度并未添加2种更具备自适应性的特征，模型仍然采用GraphSAGE模型，训练和分级策略并未采用多级分类一体化的策略，超参数设置与本方法保持一致。因为1∶50 000的比例尺中，选取与不选取的样本比例稍显平衡，且样本间特征区分度最大，所以选择在1∶50 000的比例尺下，进行了效果对比。对比结果如表4所示，常规深度学习方法在3个城市的4种选取指标都远低于本方法。由此可见，本方法相比于常规深度学习方法具有优越性。

3.3 自适应特征的消融实验

为了验证自适应特征是否能增强方法的泛化性，在1∶50 000的比例尺上进行了消融实验。实验结果表明，不使用自适应特征，南宁市和三江侗族自治县的选取效果有显著降低，且三江侗族自治县的选取效果已经不足70%。这是由于南宁市和三江侗族自治县的道路网结构有很大差别，具体体现在高等级道路的数量和类型上，因此自适应特征的效果得到充分体现；而成都市的道路网结构相对规则，各类型道路分布相对均匀，因此自适应特征带来的提升并不明显。

3.4 中心区与非中心区选取效果分析

为验证本文选取方法在城市区位差异特征保持方面的效果，从道路网长度和连通度2个维度进行定量分析。连通度通过区域道路交叉口密度（区域道路交叉口个数与区域总面积的比值）来量化。根据2020年9月自然资源部办公厅印发的《市级国土空间总体规划编制指南（试行）》，以商业化程度、发展重心偏向将城市划分为中心区域和非中心区域，计算了3个城市中心区域与非中心区域的道路网长度和连通度。

表3展示了不同比例尺下城市中心区域与非中心区域道路长度占比，可以看出，在不同比例尺的选取中，中心区域和非中心区域的道路长度占比变化差异较小，基本在同一数量级。

表3　不同比例尺下中心区与非中心区道路长度占比/%

城市	区域	1∶5 000	1∶100 000	1∶1 000 000
成都市	中心区域	8.7	9.6	10.1
	非中心区域	91.3	94.4	89.9
南宁市	中心区域	7.7	6.8	7.3
	非中心区域	92.3	93.2	82.7
三江侗族自治县	中心区域	5.8	5.5	4.7
	非中心区域	94.2	94.5	95.3

表4展示了不同比例尺下城市中心区域与非中心区域所选道路交叉口数量占比，从大比例尺到小比例尺，城市中心区域的道路交叉口数量占比逐步增高，同时结合选取的道路长度占比来看，比例尺越小，所选取的道路连通度越高，基本符合城市道路网的区位差异特征。

表4　不同比例尺下中心区与非中心区道路交叉口数量占比/%

城市	区域	1∶5 000	1∶100 000	1∶1 000 000
成都市	中心区域	14.2	14.5	23.4
	非中心区域	85.8	85.5	76.6
南宁市	中心区域	11.2	13.6	17.3
	非中心区域	88.8	86.4	82.7
三江侗族自治县	中心区域	8.6	13.5	14.4
	非中心区域	91.4	86.5	85.6

从实验结果来看，本方法在较大比例尺地图的道路选取中表现出色，取得了显著的效果。然而，在中小比例尺地图上的道路选取效果仍存在较大的改进空间。其主要原因在于，中小比例尺地图中需要选取的道路特征相对大比例尺地图中的道路特征更加复杂，现有方法未能充分挖掘道路特征的潜力。

4　结　语

本文创新性地提出一种新型道路自动选取方法。该方法融合了自适应特征与多级分类一体化的优势，同时引入图采样聚合神经网络，旨在深度挖掘节点特征之间潜藏的联系，进而显著提升道路自动选取的性能表现。实验结果表明，在针对不同比例尺地图的道路选取任务中，此方法相较于传统方法展现出更为卓越的性能。此外，通过精心设计的消融实验，进一步证实了自适应特征的引入，极大地增强了该方法在多城市道路选取任务中的通用性。这一研究成果为未来地图综合领域中道路自动选取技术的应用，提供了全新的研究思路与方向，具有重要的理论价值与实践意义。在未来工作中，将继续探索如何在中小比例尺下，更深层次挖掘道路特征的问题，以提升中小比例尺下的选取效果。

参考文献

[1] 曹炜威,张红,何晶,等. 环形放射状城市道路网的结构复杂性探讨[J]. 地理空间信息,2015,13(4):114-116

[2] Thomson R C, Richardson D E. The ‘Good Continuation’ Principle of Perceptual Organization Applied to the Generalization of Road Networks[C]//Proceedings of the ICA 19th International Cartographic Conference. Ottawa, Canada,1999:1 215-1 223

[3] Jiang B, Claramunt C. A Structural Approach to the Model Generalization of an Urban Street Network[J]. GeoInformatica,2004,8:157-171

[4] Liu X, Zhan F B, Ai T. Road Selection Based on Voronoi Diagrams and “Strokes” in Map Generalization[J]. International Journal of Applied Earth Observation and Geoinformation, 2010, 12:194-202

[5] Jiang B, Harrie L. Selection of Streets from a Network Using Self-organizing Maps[J]. Transactions in GIS, 2004, 8(3):335-350

[6] 刘凯,李进,沈婕,等. 基于BP神经网络和拓扑参数的道路网选取研究[J]. 测绘科学技术学报,2016,33(3):325-330

[7] 袁林辉,刘凯,刘佩,等. 径向基函数神经网络用于小比例尺道路网选取[J]. 测绘科学,2019,44(3):8-14

[8] 马超,熊顺,蒋丹妮. 图卷积神经网络在道路网选取中的应用[J]. 测绘科学,2022,47(12):200-205

[9] 朱余德,杨敏,晏雄锋. 利用图卷积神经网络的道路网选取方法[J]. 北京测绘,2022,36(11):1 455-1 459

[10] Scarselli F, Gori M, Tsoi A C, et al. The Graph Neural Network Model[J]. IEEE Transactions on Neural Networks, 2008,20(1):61-80

[11] Hamilton W, Ying Z, Leskovec J. Inductive Representation Learning on Large Graphs[J]. Advances in Neural Information Processing Systems,2017,30(1):1-4

[12] Zhou J, Cui G, Hu S, et al. Graph Neural Networks: A Review of Methods and Applications[J]. AI Open, 2020(1): 57-81

[13] Shannon C E. A Mathematical Theory of Communication[J]. The Bell System Technical Journal, 1948, 27(3): 379-423