基于 LA-GraphCAN 的甘肃省泥石流易发性评价

郭玲，薛晔*，孙鹏翔

太原理工大学经济与管理学院，太原 030024

第一作者，E-mail：guoling0983@link.tyut.edu.cn

* 通信作者，E-mail：xueye0412@126.com

2024-06-12 收稿；2024-08-10 修回；2024-08-14 接受

基金项目：国家自然科学基金青年科学基金项目 (41101507)

中图分类号：P642.23　　文章编号：2096-8523(2026)01-0212-13　　doi: 10.19509/j.cnki.dzkq.tb20240324

摘　要：目前对泥石流灾害易发性相关研究尚未考虑泥石流灾害的地理位置关系以及空间依赖性。本研究构建了包含 4286 个正样本点和 5912 个负样本点的甘肃省泥石流数据集，提出了一种基于 LA-GraphCAN(local augmentation graph convolutional and attention network)的泥石流易发性评价方法。首先，以样本点的经纬度投影坐标为基础，利用 KNN(K-nearest neighbors)构建最近邻图，捕捉泥石流灾害点之间的复杂地理位置关系；其次，使用 GCN(graph convolutional network)高效聚合局部邻域信息，提取关键地理和环境特征，不仅关注单个栅格所包含的信息，还深入探讨了相邻栅格之间空间结构的相互关系，从而使模型能够更精准地识别和理解样本中的局部空间特征。同时，引入 GAT(graph attention network)添加动态注意力机制，细化特征表示；再次，验证所提方法的有效性，并从不同角度对比分析；最后，对甘肃省泥石流易发性进行评价。结果表明，考虑了泥石流灾害地理位置关系的 LA-GraphCAN 的 ROC 曲线下面积(*AUC*)、准确率、精确率、召回率以及 *F*1 分数分别为 0.9868, 0.9458, 0.9436, 0.9228 和 0.9331，与主流机器学习模型 CNN(convolutional neural networks)、Decision tree 等相比最优。基于 LA-GraphCAN 评价的甘肃省泥石流极高易发区中历史泥石流灾害点数量为 4055 个，占甘肃省历史泥石流总数的 95%，与历史灾害分布基本一致。性能评估和甘肃省泥石流易发性评价结果均表明考虑泥石流灾害空间依赖性的 LA-GraphCAN 方法的评价结果更优，在泥石流易发性评价研究中有较好的适用性。

关键词：LA-GraphCAN；泥石流易发性评价；GCN；GAT；甘肃省

地质灾害在我国分布极为广泛，据中国统计年鉴统计，2000–2022 年，我国共发生地质灾害 42 万起，造成直接经济损失高达 947 亿元，其中，泥石流发生了 26000 起[1-2]。泥石流灾害不仅对生态环境造成了巨大破坏，而且对人民群众生命财产安全造成了极大威胁，已经成为制约国家经济社会发展的一个重要因素[3]。因此，加强泥石流灾害的预警，对于保障公共安全和促进可持续发展具有重要意义。

泥石流易发性评价作为泥石流灾害预警的关键手段之一，是一种定量预测研究区内泥石流易发性程度空间分布的方法。传统泥石流易发性评价方法包括定性评价方法、物理模型方法以及数理统计方法等，其中，启发式方法、专家打分法和层次分析法基于人为主观赋予指标因子权重，导致评价结果具有较强的主观性[4-5]；物理模型方法过度依赖模型调整参数，需耗费大量时间和人力成本[6-7]；常用的数理统计方法包括确定性系数法、信息量法、证据权法等，这类方法简化了评价因子之间复杂的相互作用和非线性关系[8-10]。因此，传统泥石流易发性评价方法具有评价结果主观性过强、费时费力及无法挖掘

著录信息：郭玲，薛晔，孙鹏翔. 基于 LA-GraphCAN 的甘肃省泥石流易发性评价 [J]. 地质科技通报，2026，45(1)：212-224.
GUO Ling, XUE Ye, SUN Pengxiang. Susceptibility assessment of debris flow in Gansu Province based on LA-GraphCAN[J]. Bulletin of Geological Science and Technology, 2026, 45(1): 212-224.

© Editorial Office of Bulletin of Geological Science and Technology. This is an open access article under the CC BY-NC-ND license.　https://dzkjtb.cug.edu.cn

评价因子之间复杂关系的局限性。

近年来，随着人工智能的迅速发展，众多学者将机器学习算法应用于地质灾害易发性评价。泥石流易发性评价采用的机器学习模型主要有支持向量机、逻辑回归、随机森林等[11-13]。李坤等[14]提出了一种基于少数类过采样和多粒度级联森林的泥石流易发性评价方法，该方法虽精度较高，但研究区域仅为昆明市东川区，研究样本的数量较少，未能充分验证其方法的泛化性。饶姗姗等[15]提出了一种基于 RSIV-RF 模型的泥石流易发性评价法，以四川省凉山州为研究区域，所提模型具有较强的泛化性，但精度较低。CHEN 等[16]以吉林省为研究区域，比较了 6 种 GBDT 算法在泥石流易发性评价的性能，发现 CatBoost 模型具有最高的准确率，该研究虽充分对比了 GBDT 算法在泥石流易发性评价的性能，但并未研究深度学习模型在泥石流易发性评价的性能。由于地质灾害具有很强的时空异质性，而上述方法的模型输入大多为单个栅格对应所有泥石流评价因子特征的一维特征向量，输出为单个栅格的泥石流易发性预测值，无法考虑评价因子的空间相关性[17]。

与传统机器学习模型不同，深度学习专注于构建包含多层隐层节点的模型结构，通过逐层特征变换突出特征学习的重要性，因此在表达非线性函数方面有更出色的能力。地质灾害的形成机制较为复杂，所涉及的数据量也较为庞大，深度学习模型在理论上能够更好地提取地质灾害的形成特征，高效地处理各种冗杂的数据，提供更精确的易发性评估[18-20]。神经网络模型作为机器学习领域的热点，常见的模型有 ANN（人工神经网络）、CNN（卷积神经网络）等。已有学者将神经网络模型成功地应用于滑坡易发性评价和崩塌易发性评价中，但在泥石流易发性评价中应用较少。在地质灾害易发性评价方面，神经网络模型比传统的物理模型和统计模型具有更高的精度[21-22]。邓日郎等[23]提出了一种基于 CNN 的结合信息量和间谍技术改进两步法采样策略 PU 学习框架，对广东省广州市黄埔区进行了崩塌易发性评价，使用随机森林和支持向量机对比验证，发现使用 CNN 模型的预测精度较高。ADITIAN 等[24]对比了 ANN、多元逻辑回归和频率比应用于滑坡易发性评价的效果，指出 ANN 在性能上超过另外 2 种模型。杨延晨等[25]以三峡库区万州区为研究对象，运用 CNN 进行了滑坡易发性评价，并与随机森林、支持向量机和多层感知器神经网络进行了比较，发现 CNN 在滑坡易发性评价中有更强的预测能力。但是，基于 CNN 模型的地质灾害易发性评价没有考虑灾害点间的地理位置关系，且一些人工神经网络算法有时会陷入局部最优的问题[26]。

针对现有研究和方法存在的不足，笔者以甘肃省作为研究区，构建甘肃省全省泥石流数据集，包括 4 286 处灾害点和 5 912 处非灾害点样本；以经纬度投影坐标为基础，使用 KNN（K-nearest neighbors）构建最近邻图，反映泥石流点之间的地理空间关系。将 GCN（graph convolutional network）应用于泥石流易发性的评估，通过图的结构来聚合信息，保留节点之间的全局关系，提出一种基于 LA-GraphCAN（local augmentation graph convolutional and attention network）模型的新方法，可以彻底探索泥石流中所涉及的空间依赖性。同时，LA-GraphCAN 采用局部增强技术来增强 GCN 的表达能力，集成 GAT（graph attention network），使其能够精确地识别在特定情况下哪些地理特征或环境条件更为关键，充分挖掘泥石流发生的复杂空间关系。最后与 7 种主流机器学习和深度学习模型在泥石流易发性评价中的性能进行对比，验证所提模型的有效性。

1　评价方法模型

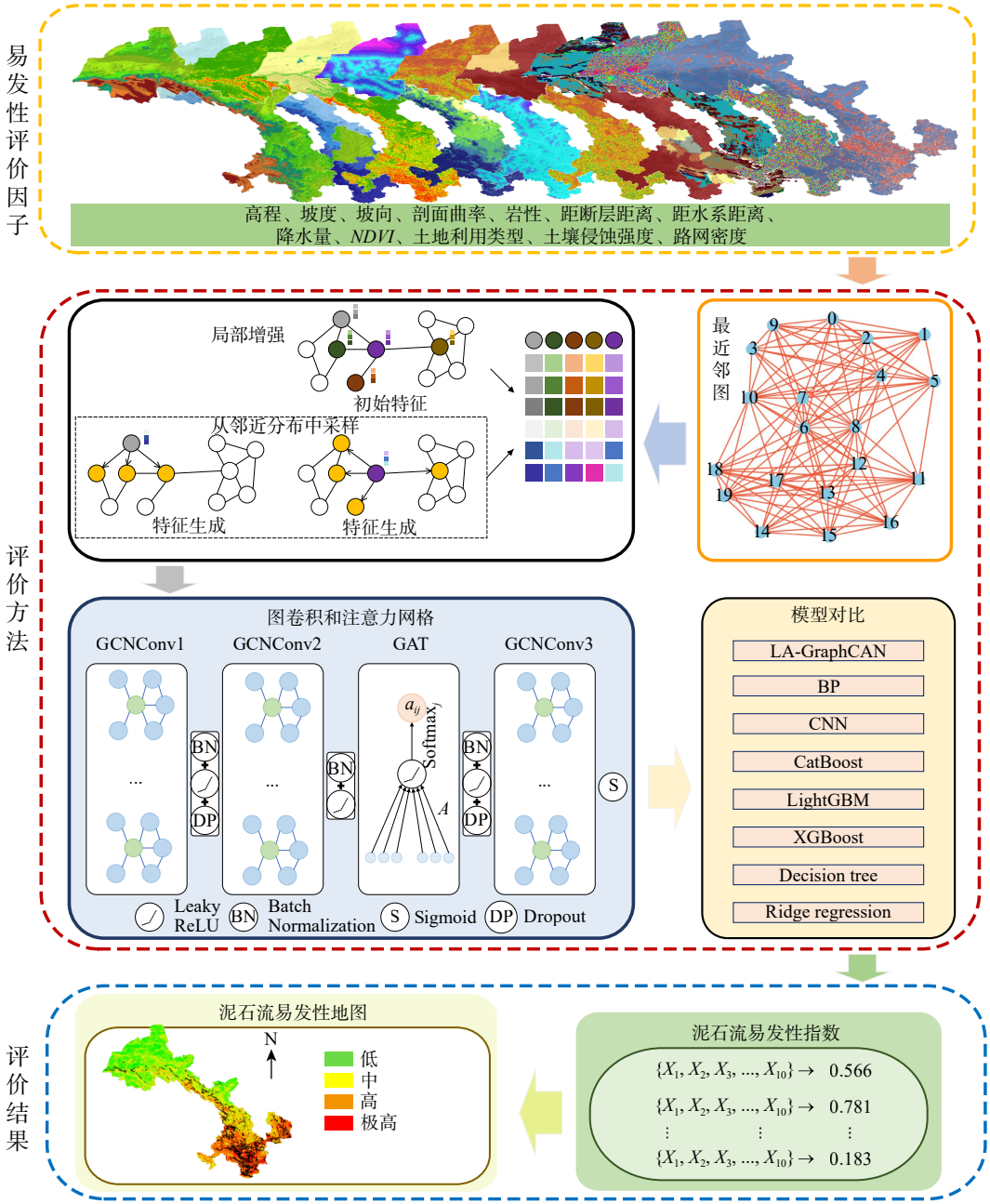
1.1　基于 LA-GraphCAN 的泥石流易发性评价流程

基于 LA-GraphCAN 的甘肃省泥石流灾害易发性评价整体框架如图 1 所示。首先，提取易发性评价因子，构建甘肃省泥石流易发性评价数据集；其次，使用经纬度投影坐标，基于 KNN 算法构建最近邻图；然后，将构建好的图数据输入 LA-GraphCAN 模型，得到泥石流易发性评价结果；最后，将该结果绘制得到泥石流易发性地图。

1.2　基于 KNN 的最近邻图构建

泥石流的发生不仅受特定地点的地形、降雨量等因素的影响，还受到周围地理环境的影响，为有效捕捉和理解泥石流发生的复杂空间关系，以经纬度投影坐标为基础，基于 KNN 构建图的邻接矩阵，如图 2 所示。图 2a 为最近邻图整体，图 2b 为最近邻图局部，节点代表特定的地理位置，而边反映了这些位置之间的相对近邻关系，该结构有助于模型理解和学习从一个地区到另一个地区可能变化的风险因素。

首先，假设泥石流数据集样本点数量为 N，通过样本点的经纬度投影坐标构建邻接矩阵，样本点即为邻接矩阵的节点，每个节点 i 包含一对经纬度投影坐标，使用 $N\times2$ 的矩阵 $\boldsymbol{X}$ 存储所有节点的经纬



NDVI. 植被归一化指数; GCNConv1. 第 1 层图卷积; GCNConv2. 第 2 层图卷积; GAT. 图注意力机制; GCNConv3. 第 3 层图卷积; Leaky ReLU. 泄漏修正线性单元; Batch Normalization. 批量标准化; Sigmoid. 神经网络中常见的激活函数; Dropout. 神经网络中常用的正则化技术; 下同

图 1　基于 LA-GraphCAN 的甘肃省泥石流易发性评价流程

Fig. 1　Debris flow susceptibility assessment process based on LA-GraphCAN

度投影坐标, 其中, 第 i 行即节点 i 的经纬度投影坐标 $(X_{i,1}, X_{i,2})$。

其次, 通过遍历计算每对节点之间的欧氏距离得到 $N\times N$ 的距离矩阵 $\boldsymbol{D}$, 对于任意 2 个节点 i 和 j, 计算两者之间的距离 D_{ij}:

$$D_{ij}=\sqrt{(X_{i,1}-X_{j,1})^2+(X_{i,2}-X_{j,2})^2}。\tag{1}$$

然后, 将计算出的距离进行升序排序, 根据排序后的距离, 选择前 K 个距离最小的对象的索引。这些索引就构成了邻居索引集合 $indices_i$。

最后, 根据最近邻索引 $indices_i$, 构建 $N\times N$ 的邻接矩阵 $\boldsymbol{A}$, 如果节点 j 是节点 i 的最近邻, 将 A_{ij} 设置为 1, 否则为 0, 如公式(2)所示。

$$\begin{cases}1, & \text{if } j\in indices_i\\ 0, & \text{else}\end{cases}。\tag{2}$$

1.3　局部增强功能（local augmentation）

为了增强 GraphCAN 的表达能力, 引入局部增强功能[27](图 3)。该方法是一个通用的框架, 可以以即插即用的方式应用于任何 GNN 模型。它根据中

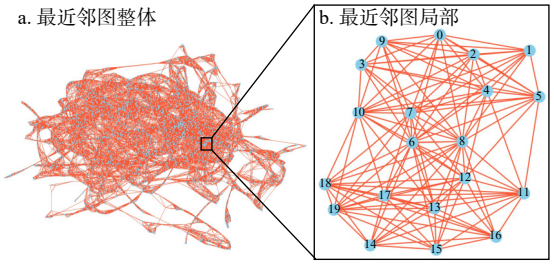


图 2　基于经纬度投影坐标的邻接矩阵

Fig. 2　Adjacency matrix based on longitude and latitude projection coordinates

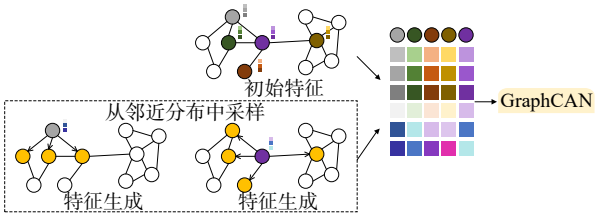


图 3　局部增强功能

Fig. 3　Local augmentation

心节点的特征来学习相邻节点特征的分布，并通过生成的特征来增强 GNN 的表达能力。

黄色圆圈表示相邻的节点。假设当地区域的分布已知，特征由局部邻域分布生成。这些原始的和生成的特性随后被用作 GraphCAN 模型的输入。使用条件变分自动编码器(CVAE)来学习给定中心节点 v 的连通邻居 $u(u\in N_v)$ 的节点特征的条件分布和潜在变量 Z 的分布。因此采用 CVAE 模型 $q_\phi(Z\mid X_u,X_v)$ 作为特征增强生成器。X_v 作为条件，因为 $X_u(u\in N_v)$ 的分布与 X_v 有关。潜在变量 Z 由先验分布 $p_\theta(Z\mid X_v)$ 产生，数据 X_u 是通过以 Z 和 X_v 为条件的生成分布 $p_\theta(X\mid X_v,Z)$ 生成的，$Z\sim p_\theta(Z\mid X_v)$，$X_u\sim p_\theta(X\mid X_v,Z_v)$。设 ϕ 表示变分参数，θ 表示生成参数，KL 为近似值与真后验值的散度。存在以下关系：

$$\begin{aligned}\lg p_\theta(X_u\mid X_v)&=\int q_\phi(Z\mid X_u,X_v)\lg\frac{p_\theta(X_u,Z\mid X_v)}{q_\phi(Z\mid X_u,X_v)}\mathrm{d}Z\\&\quad+KL(q_\phi(Z\mid X_u,X_v)\,\|\,p_\theta(Z\mid X_u,X_v))\\&\geqslant\int q_\phi(Z\mid X_u,X_v)\lg\frac{p_\theta(X_u,Z\mid X_v)}{q_\phi(Z\mid X_u,X_v)}\mathrm{d}Z\end{aligned}\tag{3}$$

且证据下界(ELBO)可定义为：

$$\begin{aligned}L(X_u,X_v;\theta,\phi)&=-KL(q_\phi(Z\mid X_u,X_v)\,\|\,p_\theta(Z\mid X_v))\\&\quad+\frac{1}{L}\sum_{l=1}^{L}\lg p_\theta\left(X_u\mid X_v,Z^{(l)}\right),\end{aligned}\tag{4}$$

式中：L 为节点 v 的邻居数；$Z^{(l)}=g_\phi(X_v,X_u,\delta^{(l)}),\delta^{(l)}\sim N(0,1)$，$\delta$ 为噪声变量，g_ϕ 为可微函数。

以上为所有节点训练 1 个 CVAE。在训练阶段，目标是使用相邻的对 (X_v,X_u) 作为输入来最大化式(4)。在生成阶段，以节点 X_v 特征作为条件，并采样一个潜在变量 Z 作为解码器的输入。

1.4　GraphCAN

笔者提出了一种基于 GraphCAN 的泥石流易发性评价模型，该模型使用基于经纬度投影坐标的图结构，能够自然地模拟泥石流样本点的空间依赖性。GraphCAN 有效结合了 GCN 和 GAT，可充分利用 GCN 的稳定性和 GAT 的灵活性。

其中，GCN 利用邻接矩阵和节点特征，通过卷积操作可高效地聚合局部邻域内的信息[28]，这种聚合机制帮助模型捕捉局部地理特征，并将其与泥石流发生的概率相关联。此外，GCN 通过学习节点的低维表示，从而在保持空间拓扑结构的同时，提取关键的地理和环境特征。

在 GCN 中引入 GAT，可实现动态注意力机制的添加，引入 GAT 后的模型不仅能静态地聚合邻近节点的信息，而且能根据不同节点和边的重要性动态地调整其关注点[29]，可更加精细地识别哪些地理特征或环境条件在特定情况下更为关键。

GraphCAN 通过 GCN 聚合邻居信息，形成稳定的特征表示，然后通过 GAT 层进一步细化这些特征表示，关注对泥石流易发性评价更为关键的特征。这种层级的处理方式使得模型不仅能捕捉到局部的地理特征，还能动态调整特征在全局预测中的贡献。GraphCAN 结构如图 4 所示。

GraphCAN 输入基于经纬度投影坐标构建的邻接矩阵 $\boldsymbol{A}$，对于 $\boldsymbol{A}$ 中第 l 层的节点特征 $H^{(l)}$，使用 GCN 更新节点特征，得到更新后的节点特征 $H^{(l+1)}$：

$$H^{(l+1)}=\sigma(\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\tilde{A}\tilde{\boldsymbol{D}}^{-\frac{1}{2}}H^{(l)}W^{(l)}),\tag{5}$$

式中：$\tilde{\boldsymbol{A}}=\boldsymbol{A}+\boldsymbol{I}$ 表示邻接矩阵 $\boldsymbol{A}$ 加自连接 $\boldsymbol{I}$；$\tilde{\boldsymbol{D}}$ 为 $\tilde{\boldsymbol{A}}$ 的度矩阵；σ 为非线性激活函数；$H^{(l)}$ 为第 l 层的节点特征；$W^{(l)}$ 为第 l 层的权重。

进一步，对 GCN 的输出进行批标准化：

$$BN(x)=\gamma\left(\frac{x-E[x]}{\sqrt{Var[x]+\varepsilon}}\right)+\beta,\tag{6}$$

式中：$BN()$ 为批标准化函数；x 为 GCN 的输出；γ，β 均为可学习参数；ε 为很小的常数，防止除数为 0；E 为均值；Var 为方差。

在第 2 层和第 3 层 GCN 之间，引入 GAT 添加动态注意力机制 $\vec{a}$，得到注意力系数 α_{ij}：

$$\alpha_{ij}=\frac{\exp(\mathrm{LeakyReLU}(\vec{a}^{\mathrm{T}}[\boldsymbol{W}\vec{h}_i\|\boldsymbol{W}\vec{h}_j]))}{\sum_{k\in K_i}\exp(\mathrm{LeakyReLU}(\vec{a}^{\mathrm{T}}[\boldsymbol{W}\vec{h}_i\|\boldsymbol{W}\vec{h}_k]))},\tag{7}$$

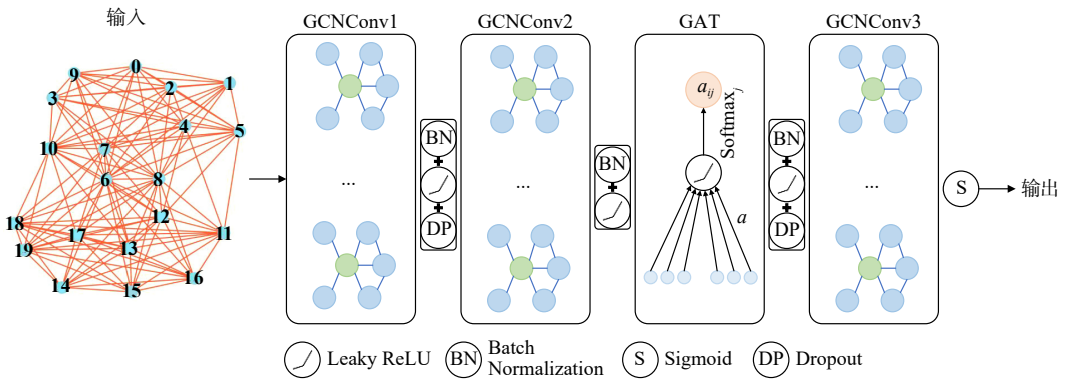


图 4　GraphCAN 结构

Fig. 4　GraphCAN structure

式中: T 为转置; ‖ 为连接操作; K_i 为第 i 个节点的邻居节点集合; a 为注意力机制, 采用单层的前馈神经网络实现; $\boldsymbol{W}$ 为权重矩阵; $\vec{h}_i$ 、$\vec{h}_j$ 分别为节点 i 和节点 j 的输入特征向量; LeakyReLU() 为激活函数。

2　研究区与数据集

2.1　研究区概况

甘肃省位于中国西北部, 地理坐标为北纬 32°11′~ 42°57′, 东经 92°13′~ 108°46′, 总面积为 425 800 km^2。甘肃省地形复杂, 山脉纵横交错, 处于黄土高原、青藏高原、内蒙古高原的夹挤之间, 西依祁连山脉, 境域形成狭长的“如意状”。甘肃省地势自西南向东北倾斜, 海拔 562 ~ 5 842 m, 相差悬殊, 高山、盆地、平川、沙漠和戈壁等地形皆有, 是山地型高原地貌。境内山谷高深, 沟壑纵横, 地质构造复杂, 地震频繁且强度高。受纬度、地形和海陆分布影响, 甘肃省的气候类型多样, 整体处于干旱半干旱区域, 少雨多风沙, 且降雨在时空分布上极不均衡, 降雨集中于 7–9 月, 局部地区暴雨多。该地区地表多见黄土、泥岩等软弱岩土体和厚层松散堆积物, 地表裸露、生态环境脆弱, 为泥石流发生提供了必要的降雨、地形和岩土条件。一些人类活动也助长了泥石流的形成, 如: 滥砍滥伐森林、开山采矿、因修建公路等工程而不合理地开挖山体等[30]。

泥石流在甘肃省的分布具有明显的地域性, 分布密度和暴发频率由南向北呈递减趋势, 主要集中在山区和高降水区, 尤其是南部和东部地区。全省沟谷型和河谷型泥石流占 80% 以上, 其余为山坡型泥石流。按甘肃省地貌分区, 大致可以划分为黄土高原区、甘南高原区、陇南山地区和河西地区。其中黄土高原区泥石流灾害最为发育, 占泥石流总数的 50%, 区域内沟谷切割深, 黄土、红层等易滑地层广泛分布。其次为陇南山地区泥石流, 占泥石流总数的 30%, 主要分布在舟曲县以下沿大断裂带的白龙江两岸[31-33]。研究区概况图及历史泥石流灾害空间分布如图 5 所示。

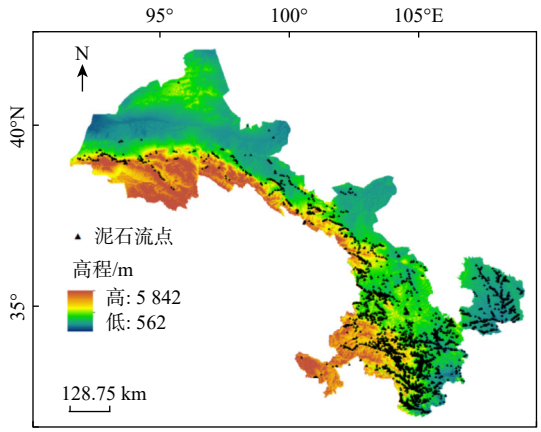


图 5　甘肃省地理位置及泥石流灾害分布

Fig. 5　Geological location and debris flow disaster distribution in Gansu Province

2.2　评价单元选取

评价单元的选取是进行地质灾害数据计算分析的重要前提, 通常采用栅格单元、斜坡单元、流域单元等对区域泥石流易发性进行评价。基于 GIS 的栅格单元技术方法成熟, 易于获取和处理各类影响因素的数据, 不需要依靠高精度的 DEM 数据做复杂的流域提取和斜坡单元提取[34-35], 笔者选用栅格单元作为评价单元, 采用 30 m×30 m 的大小进行栅格化。

2.3　评价因子选取

泥石流的形成包括 3 个基本条件: (1)陡峻的地形地貌以利于泥石流的发育运动; (2)短时间内有充沛的水源; (3)丰富的松散物质来源。泥石流的发生是多种因素共同作用的结果, 包括孕灾因子(如地形地貌、地质构造等)、诱灾因子(如: 短时间内的强降雨、地震等)。笔者根据已有研究结合甘肃省泥石流

发育特征[36-37]，综合考虑指标选取的科学性、代表性、可操作性，选取高程、坡度、坡向、剖面曲率、岩性、距断层距离、降水量、距水系距离、植被归一化指数（*NDVI*）、土地利用类型、土壤侵蚀强度、路网密度 12 个评价因子。所有评价因子均按照自然断点法进行重分类处理，如图 6 所示。

地形地貌因子在较大程度上决定了泥石流灾害的空间分布特征。高程数据受到差异性风化及人类活动强度差异的影响，不同高程的栅格单元松散物质量会有所差异。坡度的大小直接关系到斜坡的稳固程度以及地表水流的运动速率。不同坡向的斜坡因其接收阳光的程度各异，导致地表水分布、植被覆盖状况以及土壤侵蚀速率呈现出多样性。在阳光照射较为充足的区域，土壤水分蒸发更为迅速，岩石风化进程也更为显著，这些因素共同作用增加了泥石流发生的概率。

不同的岩土类型具有不同的强度、结构特征，从而形成不同厚度、不同性质的地表沉积物，影响泥石

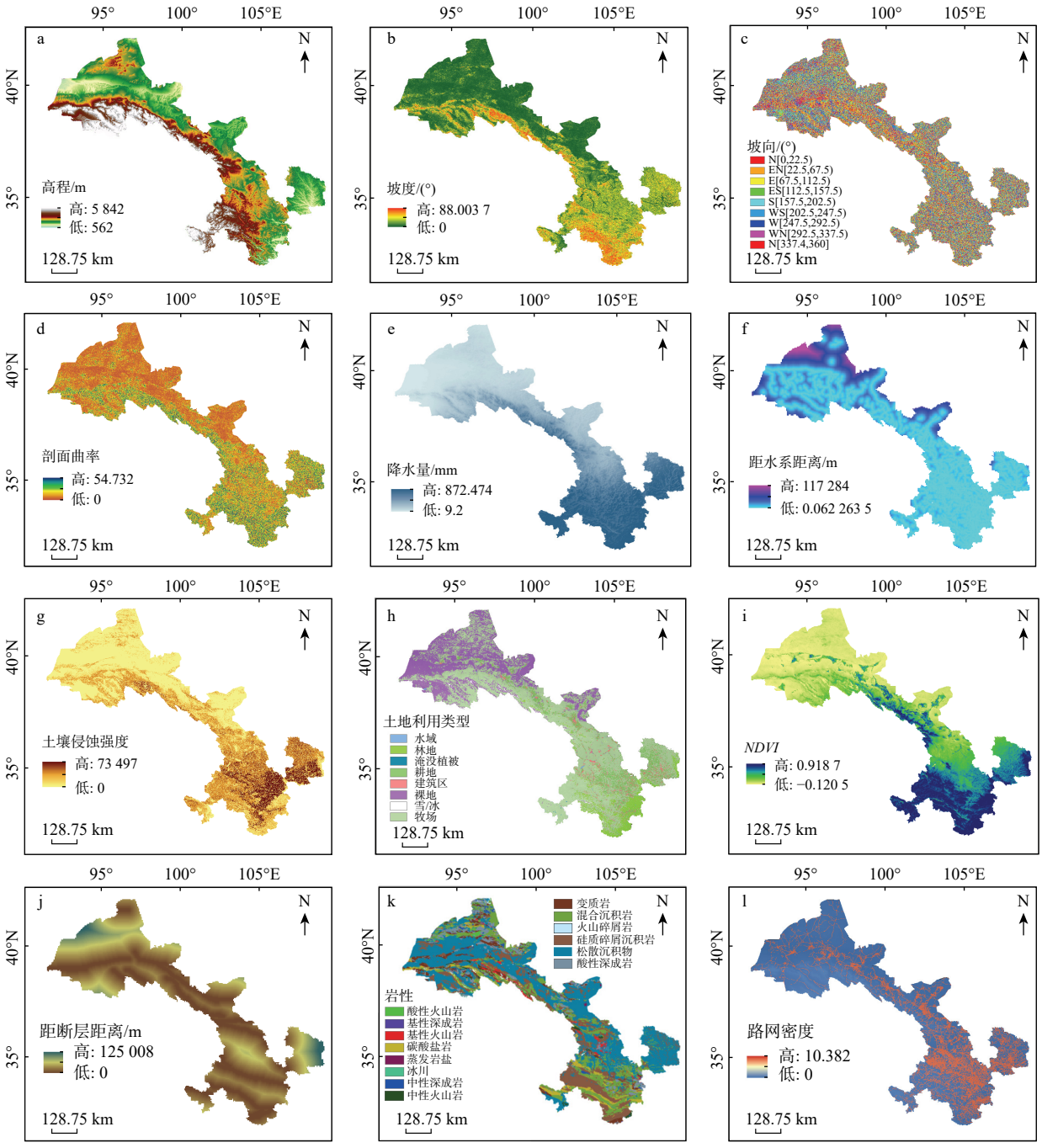


图 6　甘肃省泥石流评价因子分布

Fig. 6　Evaluation factor classification of debris flow in Gansu Province

流的物源条件。甘肃省发育的地层岩性以沉积岩为主，笔者根据研究区岩性特点，提取出主要岩组，共分为如下 14 类：酸性深成岩、酸性火山岩、基性深成岩、基性火山岩、碳酸盐岩、蒸发岩盐、冰川、中性深成岩、中性火山岩、变质岩、混合沉积岩、火山碎屑岩、硅质碎屑沉积岩、松散沉积物。

水文环境方面，降水量和河流侵蚀是泥石流灾害发生的主要诱发因素之一。降水高峰期河流、沟谷地表水暴涨暴落，是激发地质灾害、造成水土流失和坡面侵蚀最活跃的因素。

土地利用类型会影响地表径流的形成和分布，对地形稳定性有直接影响，进而影响泥石流的发生过程。土壤植被覆盖程度的高低对泥石流灾害的发生起到一定作用，一般来说，植被越茂密，受降水的作用越小，越不利于泥石流灾害的发生。*NDVI*（植被归一化指数）是遥感估算植被覆盖度研究中最常用的植被指数，其数值的大小表示着植被覆盖程度的高低。土壤侵蚀强度是泥石流重要的孕灾因子之一，高土壤侵蚀率意味着土壤大量流失，不仅减少了土壤的厚度，还破坏了土壤的结构和养分，使得土壤更容易被水流冲刷，增加了泥石流的物质来源。

地质灾害的形成，与人类工程活动也紧密相关。频繁的经济活动和开挖坡脚改变了斜坡岩土体应力状态和水文状态，坡体应力平衡遭到破坏，从而诱发地质灾害。选取路网密度作为评价因子，在公路密度大的地区，其建设开挖边坡会为泥石流的发育提供条件，在雨水的作用下土体失稳而导致泥石流易发生。

笔者把泥石流易发性研究问题转变为二分类问题。泥石流灾害点通过泥石流灾害编录库及野外调查获得，共收集 2000−2020 年甘肃省泥石流灾害点 4286 处，即 4286 处正样本记为“1”，选取非泥石流灾害点 5912 处记为“0”。使用 ArcGIS 将评价因子汇总至栅格单元，通过数据提取致点制作甘肃省评价因子数据集。

3　泥石流易发性评价与结果分析

3.1　模型训练及验证框架

泥石流易发性评价模型训练及验证流程如图 7 所示，首先，将甘肃省泥石流数据集随机按照 7∶3 进行划分，训练集占比为 70%，测试集占比为 30%；其次，使用训练集训练深度学习和机器学习模型；再次，利用测试集对模型性能进行评估；最后，对不同模型的泥石流易发性评价结果进行可视化，同时从不同维度对比模型性能。

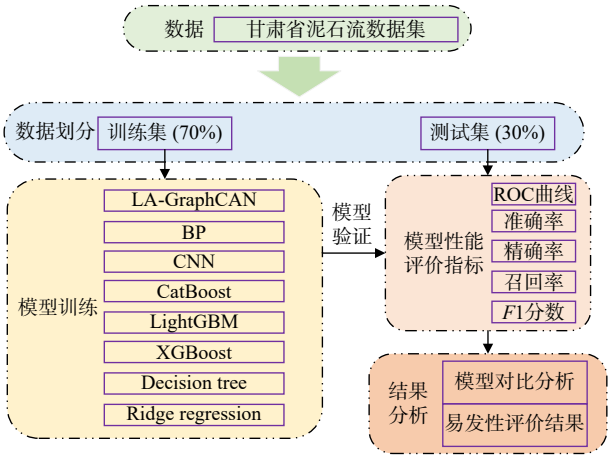


图 7　模型训练及模型验证流程

Fig. 7　Model training and validation process

3.2　模型性能评价指标

泥石流易发性评价本质为二分类任务，故采用分类任务常用评价指标准确率（*Accuracy*）、精确率（*Precision*）、召回率（*Recall*）、*F*1 分数（*F*1*score*）和 ROC（receiver operating characteristic curve）曲线评估模型的性能。模型精度评价以曲线下面积 *AUC*（area under roc curve）的大小为依据，ROC 曲线越接近左上角，ROC 曲线下的面积越大，即 *AUC* 值越接近于 1，则评价模型性能越好，精度越高[38]。公式如式（8）−式（13）所示：

$$TPR=\frac{TP}{TP+FN},\tag{8}$$

$$FPR=\frac{FP}{FP+FN},\tag{9}$$

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN},\tag{10}$$

$$Recall=\frac{TP}{TP+FN},\tag{11}$$

$$Precision=\frac{TP}{TP+FP},\tag{12}$$

$$F1score=\frac{2\times Precision\times Recall}{Precision+Recall},\tag{13}$$

式中：*TP* 为真阳性的数量；*FP* 为假阳性的数量；*TN* 为真阴性的数量；*FN* 为假阴性的数量。

3.3　参数设置

实验基于 Ubuntu 操作系统，Intel Xeon Gold 5317，GPU 为 NVIDIA A100，内存 80 G，实验代码采用 Python 编写，深度学习框架为 PyTorch1.12.1。模型参数设置如表 1 所示。

表 1　LA-GraphCAN 参数设置

Table 1　LA-GraphCAN parameter setting

参数项	参数值
最近邻数	10
GCNConv1 out_channels（第1层图卷积输出通道）	32
GCNConv2 out_channels（第2层图卷积输出通道）	64
GAT out_channels（图注意力机制输出通道）	32
GAT Heads（图注意力机制头数）	4
GCNConv3 out_channels（第3层图卷积输出通道）	1
优化器	Adam
初始化学习率	0.01
Dropout率	0.5
损失函数	二元交叉熵损失
最大迭代次数	5000

3.4　模型对比分析

为验证 LA-GraphCAN 模型的有效性，选取 7 种主流机器学习模型进行对比，包括 Decision Tree、岭回归 Ridge regression、BP 神经网络、CNN、LightGBM、XGBoost 和 CatBoost。其中，Decision tree 是基于树的机器学习模型；Ridge regression 是传统回归模型；BP 和 CNN 是深度学习模型；LightGBM、XGBoost 和 CatBoost 基于 GBDT(梯度提升决策树)的机器学习模型[39-43]。为避免这 7 种模型因参数选取不当而导致精度较低，笔者采用五折交叉验证结合网络搜索对 7 种模型选取参数。

不同模型的 4 项评价指标如表 2 所示。单一机器学习模型如 Ridge regression、Decision tree 相比其他 5 种机器学习模型表现不佳，4 项指标明显较低，均在 0.9 以下，传统的神经网络模型如 CNN 和 BP 较单一机器学习模型效果较好，但与 XGBoost、LightGBM 和 CatBoost 这类基于 GBDT 的模型相比，在 4 个评估指标上都表现不佳。GBDT 模型在 4 项指标上表现较其他模型更好，但是 GBDT 模型的召回率普遍较低，说明它们无法识别出一些实际的阳性样本，导致了大量的假阴性。相比之下，LA-GraphCAN 模型有效地挖掘了样本点的空间依赖性，不仅在召回率方面有显著改善，而且在所有指标上都表现出色。

为进一步验证评价模型的有效性，对 8 种模型绘制 ROC 曲线进行对比，结果如图 8 所示。除 Ridge regression 外，其他模型的 *AUC* 值均大于 0.9，表明其在泥石流易发性评价中均具有出色的性能。其中 LA-GraphCAN 的 *AUC* 值最高为 0.9868，说明考虑灾害点之间的地理位置关系能够有效提升易发性评价模型的性能，具有良好的应用价值。

表 2　不同模型对比

Table 2　Comparison of different models

模型	准确率	精确率	召回率	*F*1 分数
Ridge regression	0.7446	0.7227	0.6102	0.6617
Decision tree	0.8990	0.8762	0.8772	0.8767
XGBoost	0.9380	0.9360	0.9108	0.9232
LightGBM	0.9333	0.9358	0.8988	0.9169
CatBoost	0.9270	0.9277	0.8910	0.9090
CNN	0.9039	0.9151	0.8435	0.8778
BP	0.9114	0.8952	0.8874	0.8913
LA-GraphCAN	**0.9458**	**0.9436**	**0.9228**	**0.9331**

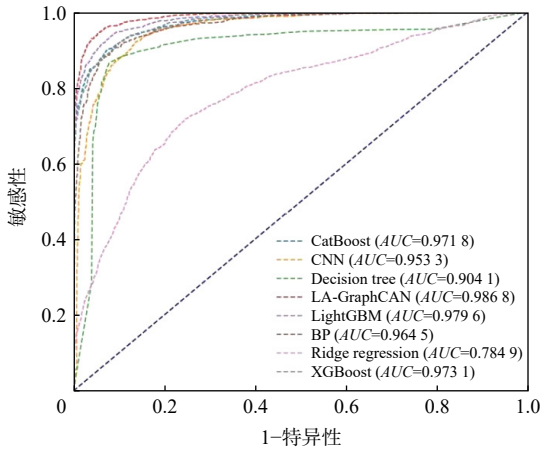


图 8　不同模型 ROC 曲线

Fig. 8　ROC curves for different models

3.5　不同图 GAT 头数对比

在 GAT 中，头数的多样性是模型对不同特征子空间探索能力的体现，每个头可以学习到不同方面的信息，而多个头的集成可以增强模型的泛化能力和鲁棒性，为确定最优的 GAT 头数，使用本研究数据集进行实验，对比不同 GAT 头数对模型性能的影响，实验结果如表 3 所示。

表 3　不同 GAT 头数对比

Table 3　Comparison of different GAT heads

GAT头数	准确率	精确率	召回率	*F*1分数
1	0.9432	0.9521	0.9123	0.9314
2	0.9425	0.9418	0.9209	0.9303
3	0.9432	0.9392	0.9184	0.9305
4	**0.9458**	**0.9436**	**0.9228**	**0.9331**
5	0.9400	0.9394	0.9127	0.9245

表 3 显示，当 GAT 头数为 1 时，模型的精度虽然达到最高，但是召回率偏低，表明模型识别正样本的能力较弱；当 GAT 头数为 4 时，模型在准确率、召回率以及 *F*1 分数上均表现最佳，表明此时模型可更

好地识别正样本，综合比较，当 GAT 头数为 4 时，LA-GraphCAN 模型在 4 个评价指标上均达到了最佳平衡，表明在本研究数据集上，当 GAT 头数设置为 4 时，模型的综合性能最优。

3.6　不同最近邻数对比

最近邻图是 LA-GraphCAN 模型的基石，而最近邻数 k 是构建图的重要参数，因此，为确定最优的 k 值，使用本研究数据集对比不同 k 值对模型性能的影响，实验结果如表 4 所示。

表 4　不同最近邻数对比

Table 4　Comparison of different nearest neighbours

最近邻数	准确率	精确率	召回率	*F*1分数
5	0.9442	0.9460	0.9162	0.9327
10	**0.9458**	**0.9436**	**0.9228**	**0.9331**
15	0.9406	0.9493	0.9101	0.9275
20	0.9332	0.9326	0.9089	0.9203

表 4 显示，当 $k = 10$ 时，模型具有最高的准确率，表明此时模型正确分类的比例最高，同时，此时的模型具有最高的召回率，表明模型识别正样本的能力最强；当 $k = 15$ 时，虽然精确度达到最高 0.9493，但是其召回率和 *F*1 分数均偏低，表明模型识别正样本较弱，局限性较大。因此，当 $k = 10$ 时，LA-GraphCAN 模型的综合性能最优，在 4 个评价指标上表现均衡且具有较高的精确度。

3.7　甘肃省泥石流易发性评价分区

基于 LA-GraphCAN 和 7 种主流模型对研究区栅格单元进行了泥石流易发性指数计算，输出得到 0 ~ 1 之间的泥石流概率值，导入 ArcGIS 中将其划分为低易发区、中易发区、高易发区和极高易发区 4 个等级，以此得到 8 种算法下的甘肃省泥石流灾害易发性评价结果，如图 9 所示选取了每类模型中的代表进行结果分析。4 种模型结果虽然存在差异，但有一些共性，对于高易发区的预测位置大致相同，绝大多数历史泥石流都包含在 4 种预测模型预测出的极高易发区。

为量化泥石流易发性评价统计结果，本研究采用 4286 个历史泥石流点作为验证数据，对不同易发性水平下的栅格数与总栅格数、泥石流密度的比例进行了统计。一个科学合理的泥石流易发性评价结果应符合随着易发性水平的增加，分级面积占比减小，泥石流密度增加[44]。在本研究中，泥石流密度是指进入各易发区的泥石流点与各易发区面积占研究区总面积的百分比。Ridge regression 结果较为发散，

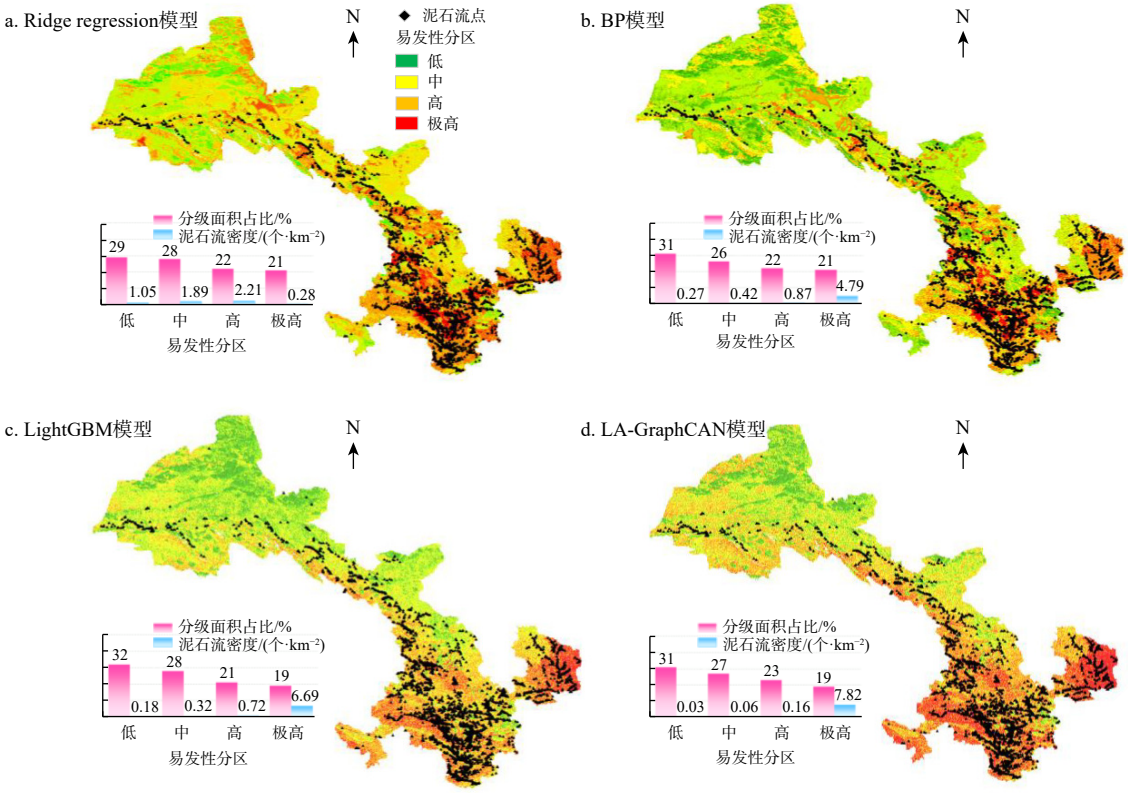


图 9　不同模型的甘肃省泥石流易发性评价

Fig. 9　Evaluation of the susceptibility assessment of debris flows in Gansu Province based on different models

历史泥石流在低、中、高易发区的数目较多；LightGBM 将部分历史泥石流点判为非泥石流点。相比之下，基于 LA-GraphCAN 的评价结果更为可靠，分级面积占比随着易发性水平的增加而减小，而泥石流密度随着易发性水平的升高而增加。在易发性极高的地区，发生泥石流事故 4055 起，泥石流密度为 7.82，在所有模型评价结果中其值最高，表明 LA-GraphCAN 在泥石流易发性评价中具有较好的适用性。

为了探讨各种因素对泥石流发生和空间分布的影响，基于 SHAP 算法对甘肃省泥石流易发性评价结果进行了解释，得到易发性因子重要性大小排序（图 10）。土地利用类型在易发性评价因子中的重要性最高，为 0.145 1，其次是路网密度为 0.143 9 和岩性为 0.143。土地利用类型会改变地表的渗透能力和水流模式，对地形稳定性有直接影响，进而影响泥石流的发生过程。岩性对泥石流的控制比降水量和高程更显著。岩土结构的差异会对土壤表面造成不同程度的大裂缝和破坏，形成大量的松散沉积物，为泥石流的发生提供了丰富的物质来源。路网密度是在评价中影响较大的另一个因素。甘肃省泥石流易发的地区主要集中在东南部，呈聚集趋势。这主要是由于该地区的人类活动密集，导致了泥石流沿着道路和河流流动聚集。此外，该地区靠近水系，雨量充沛，主要由松散的沉积岩组成，工程性质较差，为泥石流的形成提供了充足的条件。而易发性较低的地区主要位于甘肃省西北部，人口密度稀疏，年降水量较低，距离水系的距离较大。该地区人类活动的影响相对较小，降低了泥石流发生的可能性。

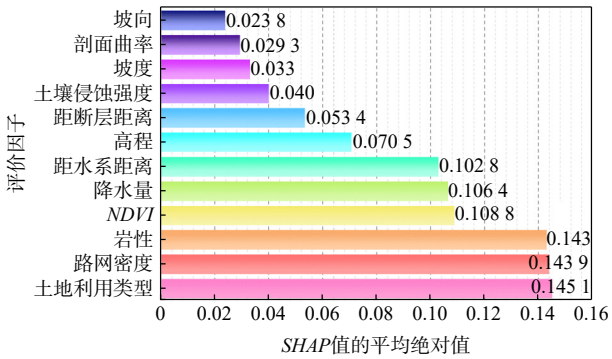


图 10　评价因子重要性排序

Fig. 10　Importance ranking of evaluation factors

4　结　论

（1）考虑到泥石流灾害有着明显的区域性和地带性分布规律，利用样本点的经纬度投影坐标构建最近邻图，捕捉泥石流灾害点之间的复杂地理位置关系，提出一种基于 LA-GraphCAN 的深度学习模型，能够高效聚合局部邻域信息，提取关键地理和环境特征，充分挖掘泥石流灾害的空间依赖性。

（2）甘肃省泥石流易发性评价结果表明，考虑灾害点地理位置关系的 LA-GraphCAN 较 7 种主流机器学习模型在 ROC 曲线下面积（*AUC*）准确率、精确度、召回率、*F*1 分数值 5 个指标上均有明显优势。

（3）基于 LA-GraphCAN 的甘肃省泥石流易发性评价分区结果显示极高易发区内的灾害点占比高达 95%，说明该模型在当地泥石流易发性评价中具有较好的适应性。易发性分区图来看东南部属于泥石流灾害高易发地区，应加强该地区的泥石流防治措施。

本研究仍存在一些不足和局限性，如：未考虑样本点的时间关联性，LA-GraphCAN 充分挖掘了样本点的空间依赖性，但并未考虑样本点的时间关联性；其次仅采取栅格单元进行评价。在后续研究中将进一步考虑样本的时间关联性，并加强基于其他评价单元方面的研究。

所有作者声明不存在利益冲突。

The authors declare that no competing interests exist.

参考文献：

[1] 史培军，刘连友. 北京师范大学灾害风险科学研究回顾与展望 [J]. 北京师范大学学报（自然科学版），2022，58（3）：458-464.
SHI P J, LIU L Y. Disaster risk science at Beijing Normal University[J]. Journal of Beijing Normal University（Natural Science），2022，58（3）：458-464.（in Chinese with English abstract）

[2] 陈丹璐，孙德亮，文海家，等. 基于不同因子筛选方法的 LightGBM-SHAP 滑坡易发性研究 [J]. 北京师范大学学报（自然科学版），2024，60（1）：148-158.
CHEN D L, SUN D L, WEN H J, et al. LightGBM-SHAP landslide susceptibility by different factor screening methods[J]. Journal of Beijing Normal University（Natural Science），2024，60（1）：148-158.（in Chinese with English abstract）

[3] 解明礼，巨能攀，赵建军，等. 区域地质灾害易发性分级方法对比分析研究 [J]. 武汉大学学报（信息科学版），2021，46（7）：1003-1014.
XIE M L, JU N P, ZHAO J J, et al. Comparative analysis on classification methods of geological disaster susceptibility assessment[J]. Geomatics and Information Science of Wuhan University，2021，46（7）：1003-1014.（in Chinese with English abstract）

[4] 赵晓燕，谈树成，李永平. 基于斜坡单元与组合赋权法的东川区地质灾害危险性评价 [J]. 云南大学学报（自然科学版），2021，43（2）：299-305.
ZHAO X Y, TAN S C, LI Y P. Risk assessment of geological hazards in Dongchuan District based on the methods of slope unit and combination weighting[J]. Journal of Yunnan University（Natural

Sciences Edition), 2021, 43(2): 299-305. (in Chinese with English abstract)

[5] 郭瑞, 马富存, 郭一兵, 等. 基于层次分析法的泥石流易发性评价 [J]. 东华理工大学学报(自然科学版), 2016, 39(4): 347-351.
GUO R, MA F C, GUO Y B, et al. Debris flow evaluation of occurrence easiness based on AHP[J]. Journal of East China University of Technology (Natural Science), 2016, 39(4): 347-351. (in Chinese with English abstract)

[6] TSAPARAS I, RAHARDJO H, TOLL D G, et al. Controlling parameters for rainfall-induced landslides[J]. Computers and Geotechnics, 2002, 29(1): 1-27.

[7] 甘建军, 罗昌泰. 中低山冲沟型泥石流运动参数及过程模拟 [J]. 自然灾害学报, 2020, 29(2): 97-110.
GAN J J, LUO C T. Runout and process simulation of gully debris flow in middle and low mountains[J]. Journal of Natural Disasters, 2020, 29(2): 97-110. (in Chinese with English abstract)

[8] 郭天颂, 张菊清, 韩煜, 等. 基于粒子群优化支持向量机的延长县滑坡易发性评价 [J]. 地质科技情报, 2019, 38(3): 236-243.
GUO T S, ZHANG J Q, HAN Y, et al. Evaluation of landslide susceptibility in Yanchang County based on particle swarm optimization-based support vector machine[J]. Geological Science and Technology Information, 2019, 38(3): 236-243.(in Chinese with English abstract)

[9] 李益敏, 杨蕾, 魏苏杭. 基于小流域单元的怒江州泥石流易发性评价 [J]. 长江流域资源与环境, 2019, 28(10): 2419-2428.
LI Y M, YANG L, WEI S H. Susceptibility assessment of debris flow in Nujiang Prefecture based on the catchment[J]. Resources and Environment in the Yangtze Basin, 2019, 28(10): 2419-2428. (in Chinese with English abstract)

[10] DOU Q, QIN S W, ZHANG Y C, et al. A method for improving controlling factors based on information fusion for debris flow susceptibility mapping: A case study in Jilin Province, China[J]. Entropy, 2019, 21(7): 695.

[11] 田春山, 刘希林, 汪佳. 基于 CF 和 Logistic 回归模型的广东省地质灾害易发性评价 [J]. 水文地质工程地质, 2016, 43(6): 154-161.
TIAN C S, LIU X L, WANG J. Geohazard susceptibility assessment based on CF model and Logistic regression models in Guangdong[J]. Hydrogeology & Engineering Geology, 2016, 43(6): 154-161. (in Chinese with English abstract)

[12] 黄发明, 陈杰, 杨阳, 等. 滑坡易发性相关致灾环境因子研究的综述与展望 [J]. 地质科技通报, 2025, 44(2): 14-37.
HUANG F M, CHEN J, YANG Y, et al. A review and prospect of disaster-causing environmental factors related to landslide susceptibility prediction[J]. Bulletin of Geological Science and Technology, 2025, 44(2): 14-37. (in Chinese with English abstract)

[13] MFONDOUM A H N, NGUET P W, SEUWUI D T, et al. Stepwise integration of analytical hierarchy process with machine learning algorithms for landslide, gully erosion and flash flood susceptibility mapping over the North-Moungo perimeter, Cameroon[J]. Geoenvironmental Disasters, 2023, 10(1): 22.

[14] 李坤, 赵俊三, 林伊琳, 等. 基于 SMOTE 和多粒度级联森林的泥石流易发性评价 [J]. 农业工程学报, 2022, 38(6): 113-121.
LI K, ZHAO J S, LIN Y L, et al. Assessment of debris flow susceptibility based on SMOTE and multi-grained cascade forest[J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(6): 113-121. (in Chinese with English abstract)

[15] 饶姗姗, 冷小鹏. 基于 RSIV-RF 模型的凉山州泥石流易发性评价 [J]. 地质科技通报, 2024, 43(1): 275-287.
RAO S S, LENG X P. Debris flow susceptibility evaluation of Liangshan Prefecture based on the RSIV-RF model[J]. Bulletin of Geological Science and Technology, 2024, 43(1): 275-287. (in Chinese with English abstract)

[16] CHEN Z L, QUAN H C, JIN R, et al. Debris flow susceptibility assessment based on boosting ensemble learning techniques: A case study in the Tumen River Basin, China[J]. Stochastic Environmental Research and Risk Assessment, 2024, 38(6): 2359-2382.

[17] 谭林, 张璐璐, 魏鑫, 等. 基于 U-Net 语义分割网络的区域滑坡易发性评价方法和跨地区泛化能力研究 [J]. 土木工程学报, 2025, 58(6): 103-116.
TAN L, ZHANG L L, WEI X, et al. Study on regional landslide susceptibility assessment method based on U-Net semantic segmentation network and its cross-generalization ability[J]. China Civil Engineering Journal, 2025, 58(6): 103-116. (in Chinese with English abstract)

[18] 黄发明, 陈彬, 毛达雄, 等. 基于自筛选深度学习的滑坡易发性预测建模及其可解释性 [J]. 地球科学, 2023, 48(5): 1696-1710.
HUANG F M, CHEN B, MAO D X, et al. Landslide susceptibility prediction modeling and interpretability based on self-screening deep learning model[J]. Earth Science, 2023, 48(5): 1696-1710. (in Chinese with English abstract)

[19] LI Y, MING D P, ZHANG L, et al. Seismic landslide susceptibility assessment using newmark displacement based on a dual-channel convolutional neural network[J]. Remote Sensing, 2024, 16(3): 566.

[20] 王毅, 方志策, 牛瑞卿, 等. 基于深度学习的滑坡灾害易发性分析 [J]. 地球信息科学学报, 2021, 23(12): 2244-2260.
WANG Y, FANG Z C, NIU R Q, et al. Landslide susceptibility analysis based on deep learning[J]. Journal of Geo-information Science, 2021, 23(12): 2244-2260. (in Chinese with English abstract)

[21] YUAN R H, LUO Y M, XU F S, et al. Mudslide susceptibility assessment based on a two-channel residual network[J]. Geomatics, Natural Hazards and Risk, 2024, 15(1): 2300804.

[22] 孔嘉旭, 庄建琦, 彭建兵, 等. 基于信息量和卷积神经网络的黄土高原滑坡易发性评价 [J]. 地球科学, 2023, 48(5): 1711-1729.
KONG J X, ZHUANG J Q, PENG J B, et al. Evaluation of landslide susceptibility in Chinese Loess Plateau based on IV-RF and IV-CNN coupling models[J]. Earth Science, 2023, 48(5): 1711-1729. (in Chinese with English abstract)

[23] 邓日朗, 张庆华, 刘伟, 等. 基于改进两步法采样策略和卷积神经网络的崩塌易发性评价 [J]. 地质科技通报, 2024, 43(2): 186-200.
DENG R L, ZHANG Q H, LIU W, et al. Collapse susceptibility evaluation based on an improved two-step sampling strategy and a convolutional neural network[J]. Bulletin of Geological Science and Technology, 2024, 43(2): 186-200. (in Chinese with English abstract)

[24] ADITIAN A, KUBOTA T, SHINOHARA Y. Comparison of GIS-

based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia[J]. Geomorphology, 2018, 318: 101-111.

[25] 杨延晨, 周超, 施佳涓. 基于卷积神经网络的区域滑坡易发性评价: 以三峡库区万州区为例 [J]. 测绘通报, 2023(11): 1-6.
YANG Y C, ZHOU C, SHI J M. Evaluation of regional landslide susceptibility based on convolutional neural network: A case study of Wanzhou District of Three Gorges Reservoir area[J]. Bulletin of Surveying and Mapping, 2023(11): 1-6. (in Chinese with English abstract)

[26] LIU P, WEI Y M, WANG Q J, et al. Research on post-earthquake landslide extraction algorithm based on improved U-net model[J]. Remote Sensing, 2020, 12(5): 894.

[27] REBETEZ J, SATIZÁBAL H, MOTA M, et al. Augmenting a convolutional neural network with local histograms: A case study in crop classification from high-resolution UAV imagery[C]//Anon. The European symposium on artificial neural networks. Ciaco: Louvain-la-Neuve, 2016: 27-29.

[28] TANG J X, ERICSON L, FOLKESSON J, et al. GCNv2: Efficient correspondence prediction for real-time SLAM[J]. IEEE Robotics and Automation Letters, 2019, 4(4): 3505-3512.

[29] BRODY S, ALON U, YAHAV E. How attentive are graph attention networks?[J]. ArXiv e-Prints, 2021, arXiv: 2105.14491.

[30] 闫满存, 王光谦. 基于 GIS 的澜沧江下游区滑坡灾害危险性分析 [J]. 地理科学, 2007, 27(3): 365-370.
YAN M C, WANG G Q. Landslide risk assessment in the lower Lancang River watershed using GIS approach[J]. Scientia Geographica Sinica, 2007, 27(3): 365-370. (in Chinese with English abstract)

[31] 刘晋文, 杨亚兵, 张磊, 等. 甘肃省 1∶5 万地质灾害风险调查成果集成 [R]. 兰州: 甘肃省地质环境监测院, 2022.
LIU J W, YANG Y B, ZHANG L, et al. Integration of 1: 50 000 geological hazard risk survey results in Gansu Province[R]. Lanzhou: Geological Environment Monitoring Institute of Gansu Province, 2022. (in Chinese)

[32] 贾强, 窦晓东, 何斌, 等. 自然灾害综合风险普查历史地质灾害调查技术报告 [R]. 兰州: 甘肃省地质环境监测院, 2021.
JIA Q, DOU X D, HE B, et al. A technical report on the survey of historical geological hazards in the general survey of the comprehensive risk of natural disasters[R]. Lanzhou: Geological Environment Monitoring Institute of Gansu Province, 2021. (in Chinese)

[33] 陈秀清, 白福, 于燕燕. 甘肃省泥石流发育特征、成因分析及其危害 [J]. 西北地质, 2014, 47(3): 205-210.
CHEN X Q, BAI F, YU Y Y. Development characteristics, causes and hazard analysis of debris flow in Gansu Province[J]. Northwestern Geology, 2014, 47(3): 205-210. (in Chinese with English abstract)

[34] 窦杰, 向子林, 许强, 等. 机器学习在滑坡智能防灾减灾中的应用与发展趋势 [J]. 地球科学, 2023, 48(5): 1657-1674.
DOU J, XIANG Z L, XU Q, et al. Application and development trend of machine learning in landslide intelligent disaster prevention and mitigation[J]. Earth Science, 2023, 48(5)1657-1674. (in Chinese with English abstract)

[35] DENG H, WU X T, ZHANG W J, et al. Slope-unit scale landslide susceptibility mapping based on the random forest model in deep valley areas[J]. Remote Sensing, 2022, 14(17): 4245.

[36] 刘国栋, 秦胜伍, 孟凡奇, 等. 基于地理信息相似度的负样本采样策略在泥石流易发性评价中的应用 [J]. 工程地质学报, 2023, 31(2): 526-537.
LIU G D, QIN S W, MENG F Q, et al. Application of geographic information similarity based absence sampling method to debris flow susceptibility mapping[J]. Journal of Engineering Geology, 2023, 31(2): 526-537. (in Chinese with English abstract)

[37] 皋子琪, 吕立群, 马超, 等. 基于 MaxEnt 模型的怒江大峡谷泥石流易发性评价和成因 [J]. 水土保持学报, 2023, 37(6): 34-41.
GAO Z Q, LYU L Q, MA C, et al. Assessment and causes of debris flow susceptibility in the Nvjiang grand canyon based on MaxEnt Model[J]. Journal of Soil and Water Conservation, 2023, 37(6): 34-41. (in Chinese with English abstract)

[38] SWETS J A. Measuring the accuracy of diagnostic systems[J]. Science, 1988, 240: 1285-1293.

[39] MULUMBA D M, LIU J K, HAO J, et al. Application of an optimized PSO-BP neural network to the assessment and prediction of underground coal mine safety risk factors[J]. Applied Sciences, 2023, 13(9): 5317.

[40] SALEHI A W, KHAN S, GUPTA G, et al. A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope[J]. Sustainability, 2023, 15(7): 5930.

[41] LIANG W Z, LUO S Z, ZHAO G Y, et al. Predicting hard rock pillar stability using GBDT, XGBoost, and LightGBM algorithms[J]. Mathematics, 2020, 8(5): 765.

[42] EL BILALI A, ABDESLAM T, AYOUB N, et al. An interpretable machine learning approach based on DNN, SVR, Extra Tree, and XGBoost models for predicting daily pan evaporation[J]. Journal of Environmental Management, 2023, 327: 116890.

[43] DAI X H, ANDANI H T, ALIZADEH A, et al. Using Gaussian Process Regression (GPR) models with the Matérn covariance function to predict the dynamic viscosity and torque of SiO_2/Ethylene glycol nanofluid: A machine learning approach[J]. Engineering Applications of Artificial Intelligence, 2023, 122: 106107.

[44] HARLOW L L. The essence of multivariate thinking: Basic themes and methods[M]. New York: Routledge, 2023.

Summary for “基于 LA-GraphCAN 的甘肃省泥石流易发性评价”

Susceptibility assessment of debris flow in Gansu Province based on LA-GraphCAN

GUO Ling，XUE Ye*，SUN Pengxiang

College of Economics and Management, Taiyuan University of Technology, Taiyuan 030024, China

First author，E-mail：guoling0983@link.tyut.edu.cn

* Corresponding author， E-mail：xueye0412@126.com

Abstract：[**Objective**] In the current studies related to the susceptibility of debris flow disasters, the geographical location relationship and spatial dependence of debris flow disasters have hitherto not been taken into account. [**Methods**] In response to this problem, this article presents a debris flow susceptibility assessment approach based on LA-GraphCAN (local augmentation graph convolutional and attention network). Firstly, a debris flow dataset for Gansu Province was constructed, encompassing 4 286 positive sample points and 5 912 negative sample points. Based on the projected coordinates of the longitude and latitude of the sample points, a nearest neighbor graph was constructed using KNN to capture the intricate geographical location relationships among debris flow disaster points. Secondly, GCN was employed to effectively aggregate local neighborhood information, extract key geographical and environmental features, and deeply explore the interrelationships of the spatial structures between adjacent grids, thereby enabling the model to more precisely identify and comprehend the local spatial characteristics within the samples. Simultaneously, GAT was introduced to incorporate a dynamic attention mechanism and refine the feature representations. Finally, the validity of the proposed method was verified and compared and analyzed from different perspectives. [**Results**] The results demonstrate that the area under the ROC curve, the accuracy rate, the precision rate, the recall rate, and the *F*1 *score* of the LA-GraphCAN model, which considers the geographical location relationship of debris flow disasters, are 0.986 8, 0.945 8, 0.943 6, 0.922 8, and 0.933 1, respectively, outperforming models such as CNN and Decision tree. [**Conclusion**] Both the performance evaluations and the assessment results of debris flow susceptibility in Gansu Province indicate that the LA-GraphCAN method, which takes into account the spatial dependence of debris flow disasters, yields superior assessment results and exhibits excellent applicability.

Key words: LA-GraphCAN； susceptibility assessment of debris flow； GCN (graph convolutional network)； GAT (graph attention network)；Gansu Province

Received：2024-06-12；Revised：2024-08-10；Accepted：2024-08-14

doi：10.19509/j.cnki.dzkq.tb20240324