

# 基于多源数据融合的工程勘察不良地质体识别方法

孙 哲

中铁第六勘察设计院集团有限公司 天津 300308

**摘 要：**针对工程勘察中不良地质体识别精度低、单一数据源存在局限等问题，本文结合空天遥感与地面勘察多源数据，构建多源数据融合识别模型。通过数据预处理、特征提取与筛选、分层融合推理等步骤，整合各数据源互补优势，弥补单一数据短板。实例验证表明，该模型识别准确率达93.2%，与现场实测吻合度92.5%，较传统方法效率提升60%以上，可精准识别滑坡、崩塌等不良地质体，为工程勘察提供可靠技术支撑。

**关键词：**多源数据融合；工程勘察；不良地质体；识别方法

**引言：**不良地质体是工程建设安全的重要隐患，精准识别其位置、规模及形态是工程勘察的核心任务。当前传统识别方法依赖单一数据源，存在勘察范围有限、精度不足、效率低下等弊端，难以适配复杂地质场景。随着数字化勘察技术发展，多源数据融合成为破解该难题的关键路径。本文立足工程勘察实际需求，构建多源数据融合识别模型，解决不良地质体识别难点，助力工程勘察数字化转型与提质增效。

## 1 相关理论与技术基础

### 1.1 工程勘察中不良地质体特征分析

**常见不良地质体类型及分布规律：**工程勘察中常见不良地质体包括滑坡、崩塌、泥石流、岩溶、采空区等。滑坡多分布于地形起伏较大、岩土体松散区域；崩塌集中在陡坡、岩性破碎地段；岩溶主要发育于碳酸盐岩分布区；采空区则与矿产开采活动密切相关，分布于矿区周边，其分布受地形、岩性、地质构造及人类活动共同影响。不良地质体核心识别指标：核心指标分为物理、力学及几何指标。物理指标含岩土体密度、含水率、孔隙度；力学指标包括抗压强度、抗剪强度、黏聚力；几何指标涵盖不良地质体规模、形态、埋深及延伸方向，这些指标是精准识别不良地质体的关键依据。不良地质体识别的难点问题：一是部分不良地质体埋深较大，地表迹象不明显；二是岩土体性质复杂，不同类型不良地质体指标交叉重叠；三是受地形、气候影响，勘察数据易出现误差，增加识别难度。

### 1.2 多源勘察数据类型及特性

**空天遥感数据及特性：**包括InSAR、机载LiDAR、倾斜摄影等，具有覆盖范围广、获取速度快、时效性强的特点，可实现大范围地质体宏观监测，但分辨率有限，难以捕捉微观地质细节。地面勘察数据及特性：涵盖地质测绘、钻探、物探等，数据精度高、针对性强，能获

取岩土体微观参数，但勘察范围有限、耗时耗力，受现场环境制约较大。多源勘察数据的异构性分析：不同类型数据在来源、格式、精度、尺度上存在差异，空天遥感数据以图像、影像为主，地面数据多为数值、文字记录，异构性易导致数据冲突，增加融合难度<sup>[1]</sup>。

### 1.3 多源数据融合核心理论

**数据融合的基本原理与分类：**基本原理是整合多源数据的互补信息，提升数据可靠性和完整性，分为数据层、特征层、决策层三级融合，分别对应不同融合深度和需求。数据融合关键技术：包括数据预处理（去噪、归一化）、特征提取（提取关键信息）、融合推理（通过算法整合数据），其中预处理是基础，直接影响融合效果。多源数据融合在工程勘察中的适配性：可弥补单一数据的局限性，结合空天遥感的宏观优势与地面勘察的微观精度，精准识别不良地质体，提升工程勘察效率和准确性，适配各类复杂勘察场景。

## 2 基于多源数据融合的工程勘察不良地质体识别模型构建

### 2.1 多源勘察数据预处理

**数据清洗与异常值剔除：**多源勘察数据受设备精度、现场环境及人为操作影响，易出现缺失值、异常值，需先开展数据清洗。空天遥感数据剔除模糊、高噪声无效影像；地面勘察数据删除钻探、物探中因仪器故障产生的异常数值。采用 $3\sigma$ 原则等统计分析法识别异常值，结合专业经验判断数据合理性，缺失值通过插值法补充，保障数据完整可靠，为后续处理筑牢基础。数据标准化与归一化处理：多源数据量纲、取值范围差异较大（如遥感灰度值与钻探强度值），直接融合易产生偏差，需进行标准化与归一化。标准化采用Z-score方法，将数据转换为均值0、标准差1的标准化数据；归一化采用min-max方法，将数据映射至[0,1]区间，消除量纲影

响,使不同类型数据具备可比性,提升融合模型运算效率与准确性<sup>[2]</sup>。数据配准与对齐:空间维度以地面勘察数据坐标系统为基准,对InSAR、LiDAR等空天遥感数据进行坐标转换与几何校正,确保空间位置精准匹配,规避空间偏移带来的融合误差;时间维度结合勘察工期与地质体变化规律,对不同时段采集的多源数据进行时间同步校正,剔除时间差异过大、与研究时段不匹配的数据,保障数据时间维度的一致性。

## 2.2 不良地质体特征提取与筛选

单一数据源特征提取方法:针对不同数据源采用针对性提取方法,空天遥感数据通过边缘检测、纹理分析等算法,提取不良地质体的宏观形态、空间分布等特征;地面勘察数据中,钻探数据提取岩土体力学参数、埋深等特征,物探数据通过反演分析提取地质体电性、磁性等物理特征,确保单一数据源的核心特征得到充分挖掘。多源数据联合特征提取:基于单一数据源特征,采用联合特征提取算法,整合不同数据源的互补信息。例如,结合LiDAR数据的地形特征与钻探数据的岩土体强度特征,构建多维度联合特征向量;通过特征融合算法,消除冗余特征,保留能精准反映不良地质体本质的联合特征,提升特征的辨识度。关键识别特征筛选与优化:采用方差分析、互信息法等特征筛选方法,剔除与不良地质体识别无关的冗余特征、相关性过高的重复特征,筛选出辨识度高、稳定性强的关键特征。同时,通过特征优化算法,对筛选后的关键特征进行权重分配,突出核心特征(如岩土体抗剪强度、不良地质体形态参数)的作用,降低次要特征的干扰,提升后续模型识别的效率和精度<sup>[3]</sup>。

## 2.3 多源数据融合识别模型设计

融合模型整体框架设计:构建“数据输入-预处理-特征提取-融合推理-结果输出”的三层融合框架,底层为数据输入层,负责导入预处理后的多源勘察数据;中间层为特征融合与推理层,完成特征提取、融合及识别推理;顶层为结果输出层,输出不良地质体识别结果及可信度评估,框架兼顾实用性和扩展性,适配不同类型不良地质体的识别需求。不同数据源的融合策略制定:针对空天遥感数据(宏观、大范围)与地面勘察数据(微观、高精度)的特点,采用分层融合策略。数据层融合实现原始数据的初步整合,特征层融合对提取的关键特征进行加权融合,决策层融合结合不同数据源的识别结果,采用投票法、贝叶斯推理等算法,输出最终识别结果,充分发挥各数据源的优势,弥补单一数据源的不足<sup>[4]</sup>。模型参数设置与优化:结合工程勘察实际场景,初始

化模型参数(如特征权重、融合系数、迭代次数),采用交叉验证法对参数进行优化。通过调整参数,降低模型过拟合、欠拟合风险,确保模型在不同勘察场景下均能保持较高的识别精度,同时优化模型运算速度,满足工程勘察的实际应用需求。

## 2.4 模型验证方法设计

验证指标选取:选取贴合不良地质体识别任务的核心验证指标,包括准确率(识别正确的不良地质体样本占总样本的比例)、召回率(实际存在的不良地质体被正确识别的比例)、F1分数(综合准确率和召回率的评价指标)及混淆矩阵,全面评估模型的识别性能,确保模型能精准识别不同类型的不良地质体,降低误判、漏判概率。验证数据集构建:选取典型工程勘察区域的多源数据作为验证数据集,涵盖滑坡、崩塌、岩溶等常见不良地质体类型,确保数据集的多样性和代表性。将预处理后的多源数据按7:3比例划分为训练集和验证集,训练集用于模型训练,验证集用于模型性能验证;同时,引入独立的测试样本,避免验证结果出现偏差,确保模型验证的客观性和可靠性。

## 3 实例验证与分析

### 3.1 实例工程概况

工程区域地理位置与地形地貌:本次实例工程选取某山区公路改扩建工程,位于我国西南地区,地理坐标介于东经103°25'-103°32'、北纬23°18'-23°25'之间。区域地形起伏较大,以中低山为主,海拔介于800-1500m,地势西高东低,坡面坡度多在25°-40°,局部存在陡坡悬崖,地貌类型主要为侵蚀构造地貌,地表植被覆盖率约65%。工程区域地质环境条件:区域地层主要为寒武系碳酸盐岩、泥页岩及第四系松散堆积物,岩层风化程度不均,局部存在破碎带。地质构造以断层、节理为主,断层走向多为北东向,节理发育密集,易引发岩土体失稳。区域地下水类型主要为岩溶裂隙水和孔隙水,地下水位埋深3-8m,对岩土体稳定性有一定影响。研究区域不良地质体分布预判:结合区域地质资料和前期踏勘,预判研究区域内不良地质体主要为滑坡、崩塌及岩溶。滑坡主要分布于公路边坡及沟谷两侧,崩塌集中在陡坡岩性破碎段,岩溶主要发育于碳酸盐岩分布区,初步预判不良地质体影响范围约1.2km<sup>2</sup>,需通过多源数据融合模型进一步精准识别。

### 3.2 多源勘察数据采集与预处理

多源数据采集方案设计与实施:采用“空天+地面”联合采集方案,空天遥感数据选用InSAR和机载LiDAR,覆盖整个研究区域,获取地形、地表形变等信息;地面

勘察数据采用地质测绘、钻探（布设钻孔12个）、物探（高密度电法、地震反射法）相结合的方式，重点采集岩土体参数和地下地质信息，采集工作历时15天，完成所有数据的采集与整理。数据预处理过程与结果分析：按照前文设计的预处理流程，先对采集的数据进行清洗，剔除无效、异常数据，补充缺失值；再进行标准化、归一化处理，消除量纲影响；最后完成空间坐标配准和时间同步对齐，将所有数据统一至同一坐标系。预处理后，数据缺失率降至1.2%以下，异常值全部剔除，数据一致性显著提升<sup>[5]</sup>。预处理后数据质量评估：采用数据完整性、准确性、一致性三个指标评估数据质量，经检验，预处理后数据完整性达98.8%，坐标配准误差 $\leq 0.5\text{m}$ ，数据准确性满足工程勘察规范要求，一致性良好，可用于后续模型应用。

### 3.3 融合模型应用与识别结果

模型在实例工程中的应用流程：将预处理后的多源数据导入构建的融合模型，依次完成单一数据源特征提取、多源数据联合特征筛选、分层融合推理，最终输出不良地质体识别结果，整个应用流程自动化运行，耗时约2小时，适配工程勘察高效需求。不良地质体识别结果呈现：模型共识别出不良地质体18处，其中滑坡7处、崩塌6处、岩溶5处，明确了各不良地质体的位置、规模及形态，其中大型滑坡2处，中型滑坡3处，小型滑坡2处，所有识别结果均以矢量图形式呈现，标注清晰、直观。识别结果的空间分布分析：识别结果显示，不良地质体主要集中在公路边坡、沟谷及断层破碎带附近，与前期预判分布趋势一致。滑坡多分布于坡度 $30^\circ$ 以上的松散堆积物坡面，崩塌集中在岩性破碎的陡坡段，岩溶主要分布于碳酸盐岩连片区域，空间分布与地形、岩性、地质构造密切相关。

### 3.4 识别结果验证与模型性能分析

现场实测验证与对比：选取8处识别出的不良地质

体进行现场实测验证，采用钻孔取样、现场监测等方式，结果显示，模型识别结果与现场实测情况吻合度达92.5%，仅1处小型岩溶因埋深较大出现轻微识别偏差，整体识别效果良好。模型识别精度与效率评估：结合验证指标，模型识别准确率达93.2%，召回率达91.7%，F1分数达92.4%，识别效率较传统勘察方法提升60%以上，有效缩短了勘察周期，降低了现场作业强度，满足工程实际应用需求。与单一数据源识别方法的对比分析：将融合模型识别结果与单一InSAR、单一钻探数据识别方法对比，单一数据源识别准确率分别为78.5%、82.3%，均低于融合模型，且存在漏判、误判情况。融合模型整合了多源数据的互补优势，有效弥补了单一数据源的局限性，识别精度和可靠性显著提升。

### 结束语

本文围绕多源数据融合在工程勘察不良地质体识别中的应用展开研究，明确多源数据特性与融合原理，构建完整识别模型并完成实例验证。结果证实该方法可有效提升识别精度与效率，弥补单一数据源局限性。本文研究仍存在不足，后续将优化特征筛选算法、拓展数据源类型，进一步提升模型在复杂地质条件下的适配性，为工程勘察安全、高效开展提供更全面的技术保障。

### 参考文献

- [1]张峰,陈传浩,樊明彬.地质勘查和深部地质钻探找矿技术研究[J].科海故事博览,2024(15):52-54.
- [2]李祥贞.矿山地质灾害预测中多源测量数据融合方法研究[J].艺术设计,2025,(8):95-97.
- [3]展贤杰.多源数据融合提升矿山地质勘测精度的策略[J].建筑技术科学,2025,(7):103-105.
- [4]曹建富.基于多源数据融合的地质调查与资源勘查精准化方法研究[J].建筑理论,2025,(5):60-63.
- [5]李飞.地质勘查中多源数据融合技术的运用分析[J].市政工程,2022,(11):136-139.