

融合图像-文本多模态探测信息的 围岩级别智能预报

甘超^{1,2,3}, 张萌^{1,2,3}, 曹卫华^{1,2,3*}

1. 中国地质大学(武汉)人工智能与自动化学院, 湖北武汉 430074
2. 复杂系统先进控制与智能自动化湖北省重点实验室, 湖北武汉 430074
3. 地球探测智能化技术教育部工程研究中心, 湖北武汉 430074

摘要: 深部地质工程地质条件复杂、隐蔽性强, 施工过程存在较多安全隐患, 准确预报围岩级别对保障高效施工至关重要。本文针对破碎岩体特征的不易界定性提出了融合图像-文本多模态探测信息的围岩级别智能预报方法。使用依据里程范围对齐的综合物探图像, 由图文跨模态匹配模型训练得到物探图像响应特征识别模型, 用于提取破碎岩体响应特征; 由目标检测模型对破碎岩体响应特征进行定位; 进而融合物探图像的文本描述特征、破碎岩体目标检测框序列特征及相关地质参数, 构建多维统一特征向量, 利用多输出随机森林模型实现探测里程范围内围岩级别的准确预报。通过文本预报结论和目标框预报结论的相互验证, 有效避免了不易界定的破碎岩体特征的误报和漏报, 实现了由多源物探图像到围岩级别等级的预报, 在实际测试中准确率达到 92%, 为深部地质工程超前智能预报工作奠定了重要基础。

关键词: 深部地质工程; 围岩级别预报; 多模态物探信息; 图像识别。

中图分类号: U452.12

收稿日期: 2026-2-1

Intelligent Prediction of Surrounding Rock Classification by Fusing Image-Text Multimodal Detection Information

Chao Gan^{1,2,3}, Meng Zhang^{1,2,3}, Weihua Cao^{1,2,3*}

1. School of Automation, China University of Geosciences (Wuhan), Wuhan 430074, China
2. Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Hubei Province, Wuhan, Hubei 430074, China
3. Engineering Research Center of Intelligent Earth Exploration Technology, Ministry of Education, Wuhan, Hubei 430074, China

Abstract: Deep geotechnical engineering involves complex, hidden geology and high risks, making accurate rock classification critical. This study presents an intelligent prediction method that fuses image-text multimodal geophysical data via mileage-aligned survey images. A cross-modal matching model extracts fractured-rock features, localized by an object detector. Text descriptors, bounding-box sequences, and geological parameters are combined into a unified vector; a multi-output random forest predicts rock classes along mileage. Mutual correction between text and detection reduces false alarms and missed detections. Field tests show 92% accuracy, supporting intelligent advance forecasting.

基金项目: 国家自然科学基金重点项目 (No. 62333019)

作者简介: 甘超 (1990-), 男, 副教授, 主要从事复杂系统建模与优化控制的研究工作. E-mail: ganchao@cug.edu.cn

***通讯作者:** 曹卫华 (1972-), E-mail: weihuacao@cug.edu.cn

Key words: Deep Geological Engineering; Surrounding Rock Classification; Multimodal Geophysical Information; Image Recognition.

0 引言

随着我国基础设施建设迈入高质量发展阶段，深部地质掘进作为地下工程建设的核心环节，面临复杂地质构造和施工安全的双重挑战（Zou *et al.*, 2025; Sun *et al.*, 2025）。围岩级别是指隧道围岩的整体工程地质条件等级，通常根据岩体的物理性质、力学特性以及水文地质条件进行划分。围岩稳定性作为施工成败的关键因素，准确预报围岩级别对于开挖工法的动态调整，以及工程安全性、经济性和施工质量的平衡均至关重要。

围岩级别预报研究主要可分为运用单一探测信息的围岩级别预报和基于多源探测信息融合的围岩级别预报两类。

在运用单一探测信息的围岩级别预报方法中，钻探法常被用于围岩级别的量化判定。传统经验性方法如岩石质量分级法（Rock Mass Rating, RMR）（Sabri *et al.*, 2025）和岩石质量分级法（Barton Q-System, BQ）（Li *et al.*, 2025）基于钻探过程中获取的岩石物理参数进行围岩质量评价。随着人工智能技术的发展，支持向量机等机器学习算法也被引入围岩级别的智能分类与预测，实现了对复杂地质数据的更深层次挖掘（Zhang *et al.*, 2025）。然而，钻探作为点状探测手段，存在“一孔之见”的局限，难以满足掌子面前方大范围连续预报围岩级别的需求。

相比之下，物理探测技术凭借远距离、大范围的探测能力成为远距离围岩预报的重要途径（Gan *et al.*, 2025）。常用的单一物探方法包括地震波探测（Tunnel Seismic Prediction, TSP）、瞬变电磁法（Transient Electromagnetic Method, TEM）和地质雷达（Ground Penetrating Radar, GPR）。TSP 通过波速变化反映岩体完整性，能识别破碎带及其空间分布（王汪汪等, 2021）；TEM 对含水体和软弱岩层高度敏感，可通过视电阻率判定地下水状况（张旭杰等, 2023）；GPR 则以高分辨率和对破碎岩体的灵敏响应，被广泛应用于岩体结构及节理裂隙发育带识别（林志等, 2021; 高树全等, 2021）。尽管单一物探技术各有优势，但均难以全面反映复杂深部地质环境中的多样化结构特征。

为提升围岩预报的准确性和可靠性，多源探测信息融合方法被广泛研究和应用。通过整合如 TSP 与 TEM 等多种探测手段的联合三维成像，实现对含水构造和破碎带的精准定位（张硕等, 2024）；结合

“长短距离探测”的综合信息，有效修正设计阶段围岩级别，防止因岩体完整性判断失误引发的工程灾害（龙贵云等, 2024）。然而，当前对多源探测信息的解译工作仍主要依赖人工经验，定性分析异常响应特征，导致预报过程主观性强，标准化和自动化程度不足，难以满足现代深部地质工程对高效、精准预报的需求（Zhao *et al.*, 2024）。

为突破传统人工解译的瓶颈，基于深度学习的探测图像智能解译技术迅速发展，主要通过图像识别和目标检测方法实现对地质缺陷的自动识别，如裂缝、空洞、破碎带等反映岩体完整性受损的异常体（Liu *et al.*, 2025; 房振华等, 2023; 覃紫馨等, 2023）。具体技术包括卷积神经网络（CNN）及其衍生的目标检测算法，如 YOLO 系列和 Faster R-CNN，这些算法能够从物探图像中高效定位和识别地质异常特征。例如，轻量化注意力机制集成的 TLAD-YOLO 模型实现了衬砌背后空洞与不密实区的快速检测（Zheng *et al.*, 2025）；IR-TAG 算法进一步提升了断层破碎带和富水带的识别精度（蒋源等, 2023）。此外，结合图像分割的 U-Net 网络与多层感知机（MLP）能够对掌子面节理网络进行精细分割，并自动预测地质强度指标（Dong *et al.*, 2025）。尽管上述方法在探测图像局部纹理特征提取方面表现良好，但在围岩级别预报中仍有局限：破碎岩体边界模糊、形态不规则，目标框标注依赖人工经验；物探报告中专家多采用“同相轴连续性较差”“不连续”等全局性表述，从整体语义到局部目标框标注主观性较强，破碎带响应又常与噪声混杂，使单一视觉模型易产生误报、漏报。

为此，本文引入融合文本语义的视觉-语言多模态模型，将专家报告中的破碎岩体描述编码，并与物探图像在统一特征空间对齐，在图像之外增加语义约束，缓解目标框标注的不确定性。通过图像-文本双模态的信息耦合与相互校正，可在标注主观性强、信号混杂时减少误差，提升破碎岩体识别及围岩级别预报的可靠性。CLIP 模型基于大规模图文预训练实现视觉特征与文本描述对齐（Yu *et al.*, 2025），为该协同机制提供基础。该技术虽在深部地质领域起步，但在相关领域已有成功应用，如 YOLO-CLIP 区分地铁视觉相似动作（An *et al.*, 2025）、光伏缺陷小样本高精度识别（Saeed *et al.*, 2024）、CLIP-YOLO 的零样本推理（Li *et al.*, 2025）。此外，GeoPredict-LLM 实现了地质图像与文本融合（Xu *et al.*, 2025）。

al.,2025),验证了多模态融合优势。结合目标检测的局部定位和视觉-语言模型的全局语义理解,通过图文融合可提升破碎岩体识别准确性,建立物探图像特征与围岩级别语义的映射,助力实现更精准的围岩级别智能预报。

基于此,本文提出一种融合图像-文本多模态探测信息的围岩级别智能预报方法。由图文跨模态匹配模型训练得到物探图像响应特征识别模型,用于提取破碎岩体响应特征;由目标检测模型对破碎岩体响应特征进行定位;进而融合物探图像的文本描述特征、破碎岩体目标检测框序列特征及地质参数,构建多维统一特征向量,利用多输出随机森林实现对不同里程位置围岩级别的精准预报。

1 概述

本文针对深部地质工程中物探图像丰富但破碎岩体特征标注困难的问题,提出了一种融合图像-文

本多模态探测信息的围岩级别智能预报方法,整体框架如图1所示。方法利用实际施工产生的物探报告,提取图像、响应特征专家文本描述和施作面情况等地质信息。

该方法包括三大核心环节:一是构建图像及其响应特征的文本标签和目标框标签数据集,训练图像-文本跨模态匹配模型以实现响应特征的文本提取,涵盖 GPR 波列图波形、TSP 纵波波速和 TEM 视电阻率特征;二是基于破碎岩体目标检测模型定位各类破碎特征区域,如 GPR 同相轴错断、TSP 波速下降及 TEM 低阻区;三是融合多模态信息进行岩体完整性预测,并结合岩性、风化程度及岩石抗压强度计算岩石硬度,最终完成围岩级别的综合预报。

该框架有效利用图像-文本信息,减轻了数据标注负担,降低了因标注不准确引起的误报与漏报,同时提升了预报结果的可靠性和可解释性。

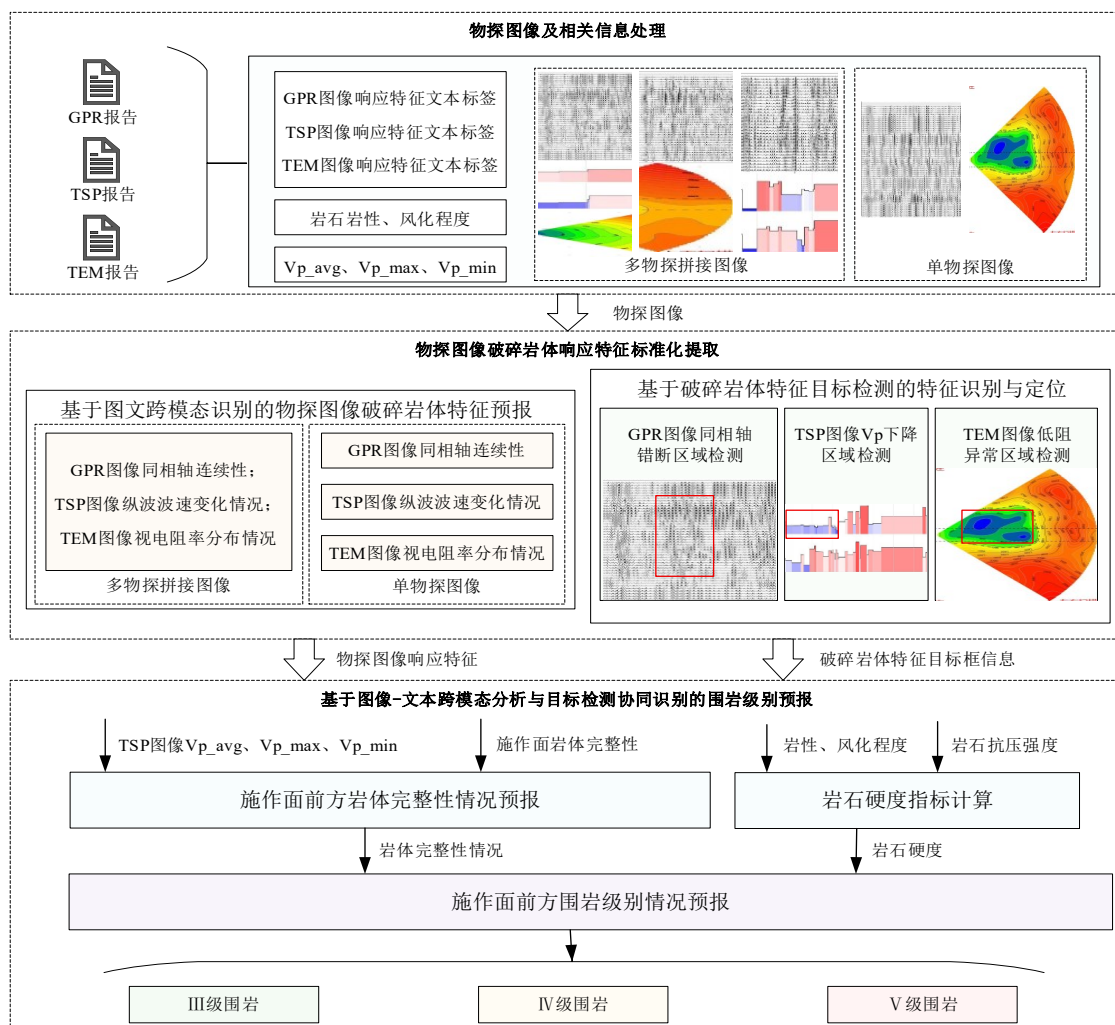


图1 方法总体框架

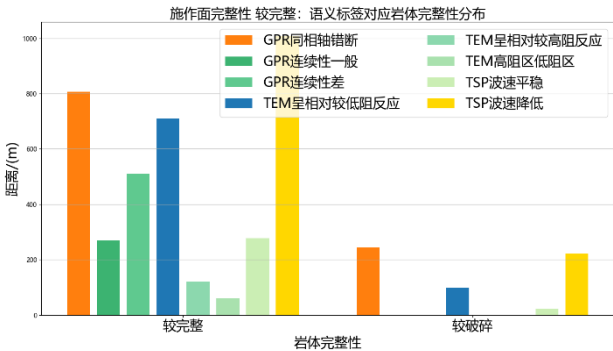
Fig. 1 Overall framework of the method

2 方法

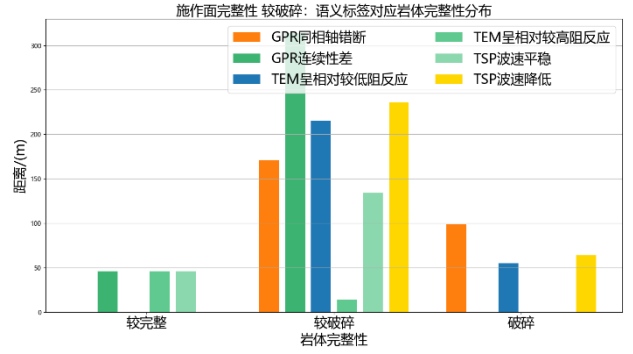
本章将阐述融合图像-文本多模态探测信息的围岩级别智能预报方法，该方法利用了各物探报告中的物探图像、专家文本表述、岩性、风化程度等关键地质信息。首先分析了图像特征标签与实际围岩级别间的相关性；然后结合图像-文本跨模态匹配破碎岩体特征提取模型和特征定位模型，得到更准确的破碎岩体响应特征信息，从而基于岩体完整性预报模型预报探测里程内的岩体完整性情况；进一步根据岩性和风化程度计算岩石硬度，最终结合岩体完整性情况和岩石硬度得到围岩级别情况。在接下来的小节中，将依次对上述各模块的详细技术路线进行论述。

2.1 图像标签类别与围岩级别相关性分析

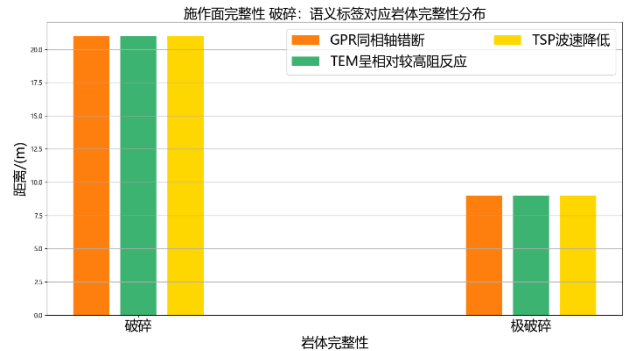
本文从各物探报告中提取的文本标签包括：同相轴错断、同相轴连续性一般、呈相对低阻反应、波速降低、波速平稳等，来自报告中专家对物探图像的分析。提取各报告中物探图像和文本标签后，构建了图像-文本一一对应的数据集。为证明方法的可靠性，对多源物探图像文本标签与实际岩体完整性、围岩级别的相关性展开分析，结果表明当文本标签出现破碎岩体响应特征的相关表述时，岩体完整性及围岩级别有变差趋势。如图 2 所示，图像横轴表示实际开挖揭示的岩体完整性情况，纵轴表示各岩体完整性覆盖的范围（单位：米），图中橙色、黄色、蓝色柱状条分别表示“同相轴错断”“波速降低”“呈相对低阻反应”破碎岩体响应特征文本标签，可以看到，在不同施作面情况下，若图像对应的文本标签含破碎岩体响应特征则实际开挖揭示岩体完整性较施作面存在变差的趋势。证明了融合文本信息辅助图像识别的可靠性。



(a)施作面较完整时文本标签对应岩体完整分布情况



(b)施作面较破碎时文本标签对应岩体完整分布情况



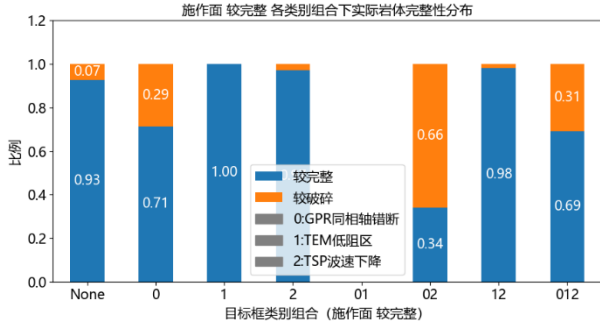
(c)施作面破碎时文本标签对应岩体完整分布情况

图 2 图像文本标签与岩体完整性分布相关性

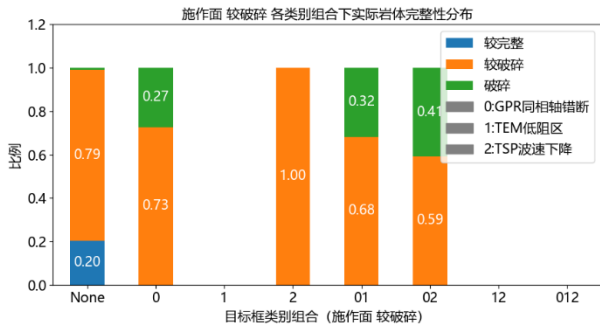
Fig. 2 Correlation Between Image-Text Labels and Rock Integrity Distribution

本文针对物探图像标注的破碎岩体目标框信息主要包括：同相轴错断具体区域、纵波波速降低区域、呈相对低阻反应的区域位置，标注操作主要依据各物探报告中专家对于物探图像的分析，标注信息包括目标区域的类别、坐标位置信息。为进一步分析破碎岩体目标框信息与实际岩体完整性的相关性，本文分析了各多源物探图像中三种类别目标框的出现情况和对应实际岩体完整性的分布情况，结果表明当图像中出现“同相轴错断”“波速降低”“呈相对低阻反应”中的一种或多种响应特征时，实际开挖揭示的岩体完整性存在变差的趋势，证明了基于图像响应特征实现岩体完整性预报的可靠性。

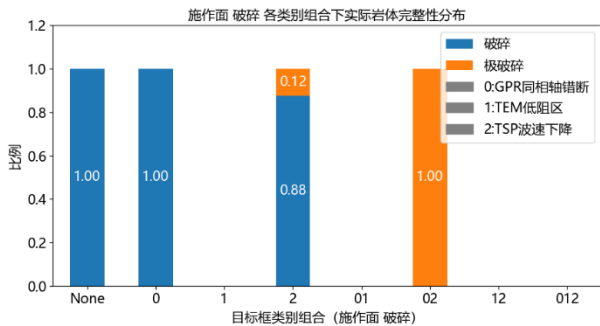
如图 3 所示，分别展示了不同施作面情况下目标框与实际岩体完整性的相关性情况，图像横轴表示探测里程范围内被目标框覆盖的情况，None 表示该里程未被任何目标框覆盖，01、02、12 表示同时出现两种特征目标框，012 表示同时出现三种特征目标框；图像纵轴表示经统计，被特征目标框覆盖的里程实际开挖揭示的岩体完整性情况概率。



(a)施作面较完整时目标框类别对应岩体完整分布情况



(b)施作面较完整时目标框类别对应岩体完整分布情况



(c)施作面较完整时目标框类别对应岩体完整分布情况

图3 图像目标框类别与岩体完整性分布相关性

Fig. 3 Correlation Between Image Bounding Box Categories and Rock Integrity Distribution

从图3呈现的统计结果来看，物探图像中破碎岩体特征目标框覆盖的区域内，岩体完整性较施作面存在变差的趋势，特别是“同相轴错断”“纵波波速下降”区域内岩体完整性变差的趋势更加显著。证明了依据物探图像破碎岩体响应特征类别及位置信息实现岩体完整性预报的可靠性。准确的岩体完整性是预报围岩级别情况的必要因素。

2.2 基于破碎岩体特征目标检测的特征识别与定位

本研究基于自建的物探图像破碎岩体特征检测

数据集，采用目标检测深度学习算法，实现对物探图像中岩体破碎区域的自动识别。具体利用经过专门训练的目标检测模型，高效定位并分类三类关键破碎岩体响应特征：地质雷达波列图中同相轴错断区域（类别0）、地震纵波波速下降区域（类别2）和瞬变电磁低视电阻率区域（类别1）。分析对象包括单一物探图像及根据里程拼接的综合物探图像。

针对地质雷达波列图中破碎岩体特征隐蔽且难以人工识别的问题，采用YOLO模型凭借其单阶段检测结构和强大的全局特征提取能力，对小尺度且分布不规则的异常区域实现高效检测。相比传统两阶段检测器，该模型兼具高精度与实时推理效率，适合现场快速响应。其多尺度特征融合和先验框设计，有效提升了对不同尺寸和形态岩体破碎区域的适应性，增强了模型泛化能力，尤其适用于复杂纹理的深部地质环境。

为提升模型泛化和鲁棒性，减轻样本量有限和类别不平衡带来的影响，本研究对原始图像及标签进行了系统的数据增强。采用Albumentations库的多种几何与色彩变换（如随机翻转、亮度对比调节、平移缩放、尺寸重采样等），并确保目标边界框同步变换，保持标注与图像内容一致。数据增强使数据量扩大约20倍，显著提升了训练稳定性与识别能力，减少过拟合风险。

模型推理阶段，输入单张物探图像，输出多个检测框，包含类别、空间位置及置信度。位置以像素坐标表示，后归一化至[0,1]区间，公式如(1)所示，保证不同分辨率图像间尺度一致，便于多模态数据融合分析。

$$\hat{x} = \frac{x_c}{W_{img}}, \hat{y} = \frac{y_c}{H_{img}}, \hat{w} = \frac{w_c}{W_{img}}, \hat{h} = \frac{h_c}{H_{img}} \quad (1)$$

其中， (x_c, y_c) 为检测框中心在原始图像坐标系下的像素坐标， w_c 和 h_c 分别为检测框在水平方向和垂直方向上的像素宽度和高度； W_{img} 和 H_{img} 分别表示该物探图像的宽和高（单位均为像素）； $\hat{x}, \hat{y}, \hat{w}, \hat{h}$ 为对应的归一化中心坐标与宽高。

此举保证了检测结果在不同图像分辨率或尺寸下的尺度一致性，有利于多图像、多模态数据融合分析中的空间特征对齐和统一处理。

所采用的目标检测模型针对本项目的特殊任务，通过标注和训练自定义数据集，在检测准确率和召回率之间取得良好平衡，能有效识别并定位典型的破碎岩体响应特征区域。该方法显著提高了物探图像中破碎岩体特征的自动提取效率，避免了传统人

工判读的主观性和时间成本，提高了数据处理的自动化和智能化水平。

综上，本研究基于先进深度学习检测技术，实现了物探图像中破碎岩体特征的自动定位与分类，为岩体完整性及围岩级别预报提供了坚实的数据支撑。

2.3 基于图像-文本跨模态识别的物探图像破碎岩体特征预报

为减少上述破碎岩体特征目标检测模型误报、漏报对最终预报结果的影响，本文引入了图像-文本跨模态匹配模型，依据物探报告中的物探图像及其对应的特征分析文本内容，构建破碎岩体特征预报模型。本研究采用 CLIP 预训练跨模态模型，该模型基于对比学习框架，能够在统一的语义空间中同时编码和匹配图像与文本信息。之所以选用该模型，一方面在于其具备良好的图文对齐能力和开放词汇表特性，适合处理地质专家以短语或句子形式给出的多样化响应特征描述；另一方面，相较于其他语义大模型，CLIP 对计算资源的要求相对较低，更易在中等配置的 Windows 平台上完成基于小规模专用数据集的微调与部署，便于与本地数据库及现有算法流程集成调用。结合本研究以短语级文本标签为主的标注形式，CLIP 在模型容量、计算开销与表达能力之间取得了较为合适的平衡，能够满足隧道物探场景下破碎岩体响应特征跨模态匹配的实验需求。

该模型包含一对深度编码器，分别用于视觉模态和文本模态的信息提取。视觉编码器选用了 ResNet-50 架构作为视觉特征提取的主干网络，通过卷积神经网络对输入的物探图像进行特征抽取；文本编码器则利用变换器结构将文本描述转化为向量表示，采用基于 Transformer 的架构，文本描述 T 首先经过分词处理，随后输入到包含多头自注意力机制的 Transformer 层中，将文本序列转化为与图像特征同维度的向量表示。对于给定的图像 I 与对应文本 T ，模型分别计算其特征向量：

$$V_I = f_I(I) \in R^d, V_T = f_T(T) \in R^d \quad (2)$$

其中 d 表示共享语义空间的维度。为了实现图像与文本特征的对齐，模型采用了基于余弦相似度的对比损失，定义两者相似度为：

$$s(V_I, V_T) = \frac{V_I V_T}{\|V_I\| \|V_T\|} \quad (3)$$

在训练过程中，模型最大化正确匹配的图文对相似度，最小化不匹配对的相似度，以提高跨模态匹配的准确性。具体的损失函数可表示为：

$$L_{I-T} = -\frac{1}{N} \sum_{i=1}^N \left[\log \frac{\exp\left(\frac{s(V_{I_i}, V_{T_i})}{\tau}\right)}{\sum_{j=1}^N \exp\left(\frac{s(V_{I_i}, V_{T_j})}{\tau}\right)} \right] \quad (4.1)$$

$$L_{T-I} = -\frac{1}{N} \sum_{i=1}^N \left[\log \frac{\exp\left(\frac{s(V_{I_i}, V_{T_i})}{\tau}\right)}{\sum_{j=1}^N \exp\left(\frac{s(V_{I_j}, V_{T_i})}{\tau}\right)} \right] \quad (4.2)$$

$$L = L_{I-T} + L_{T-I} \quad (4.3)$$

其中 N 为批次大小， I_i 表示第 i 张物探图像， T_i 为与之对应的特征分析文本， V_{I_i}, V_{T_i} 分别为图像与文本在共享语义空间中的特征向量表示， τ 是温度参数，用以调整相似度分布的平滑度。 L_{I-T} 表示以图像为查询、在当前批次中所有文本中检索匹配文本的对比损失； L_{T-I} 则表示以文本为查询、在所有图像中检索匹配图像的对比损失。两个方向的损失在式 (4.3) 中相加，构成最终的双向图文对比学习目标，从而同时提升图像检索文本与文本检索图像的跨模态匹配能力。

本文采用图像-句子级的跨模态对齐方式：以物探图像作为视觉输入，以物探报告中专家给出的相应图像响应特征文本表述作为文本输入，两者构成一对“图像-文本特征”的样本对，通过对比学习优化图像与文本在统一特征空间中的相似度分布，使模型能够稳定区分不同类型的破碎岩体响应模式。本研究所使用的物探数据涵盖深部地质工程中三类关键物探图像：地质雷达波列图、地震波纵波波速图及瞬变电磁视电阻率图。这些图像分别反映了岩体的不同物理属性，本文以拼接组合形式构成多模态输入。该跨模态识别模型能够依据输入中实际存在的物探图像组合状态，输出相应的地质响应特征描述。该方法不仅有效整合了复杂多模态信息，提升了异常地质特征识别的自动化水平，也为围岩级别预报提供另一维度支撑，减少单一预报模型的误报和漏报，具有显著的工程应用潜力和推广价值。

2.4 岩石硬度指标分析与计算

围岩级别预报通常依赖岩体完整性和岩石硬度指标，后者是围岩稳定性和支护设计的重要参数。本研究结合项目需求，提出了一种基于岩石单轴抗压强度、风化等级和岩性类型的岩石硬度分级方法，实现硬度的量化评价。

具体而言，硬度判断首先以岩石单轴抗压强度 R_c (MPa) 为主要参考指标。通过预设的四个阈值 t_1, t_2, t_3, t_4 ，将硬度划分为五档，采用半开区间形式

定义各档范围，即：

$$(-\infty, t_1), [t_1, t_2), [t_2, t_3), [t_3, t_4), [t_4, +\infty) \quad (5)$$

其中，对应从极软岩至极硬岩的等级划分，确保划分连续明确。

其次，结合风化等级（未风化到全风化）对硬度进行调整：轻微风化不变，中重度风化降低硬度一到两个档次，全风化则归为极软岩。此外，岩性按矿物成分和形成条件分组，不同组别根据力学特征赋予偏置调整，使硬度更贴合实际物理性质。最后引入整体校正偏置，保证结果合理。

综上，硬度评定依次输入单轴抗压强度、风化等级和岩性，通过分档、风化修正、岩性偏置和整体偏置，输出五级岩石硬度（极软岩至极硬岩）。该方法结合多参数综合调整，科学且实用，为深部地质工程施工安全和支护设计提供技术支持。

2.5 基于图像-文本跨模态分析与目标检测协同识别的围岩级别预报

本研究针对岩体完整性的分级预测问题，提出并实现了一种基于多源信息融合的多输出多分类模型。该模型充分整合了多模态输入数据，包括由预训练跨模态模型生成的响应式文本特征、施作面岩体完整性指标、纵波波速参数以及基于目标检测技术提取的地质异常目标框信息，构建兼具语义表达与空间层级结构的高维特征表示，进而实现对沿线 30 米范围内、以每米为单位的岩体完整性等级分类预测。

在数据预处理环节，文本描述信息作为物探图像相关地质异常语义的表达，经关键词匹配与粗粒度语义强度编码形成数值特征向量 $x \in R^d$ ，其中特征维度 d 涵盖语义文本编码、施作面完整性指标的重复列增强、纵波波速统计特征、基于目标检测的空间序列特征及其交互项。

施作面岩体完整性指标作为反映岩体质量的关键参数，被作为特征重点列进行了重复扩展，增强其在模型聚合过程中的作用。纵波波速数据包含最大值 V_p^{max} 、平均值 V_p^{avg} 、和最小值 V_p^{min} 三项指标，通过计算差异与范围值，体现岩石弹性波传播的物理特性，如式（6）：

$$\begin{cases} vp_{uprange} = |V_p^{max} - V_p^{avg}|, \\ vp_{downrange} = |V_p^{min} - V_p^{avg}|, \\ vp_{range} = |V_p^{max} - V_p^{min}|. \end{cases} \quad (6)$$

其中， $vp_{uprange}$ 为纵波波速最大值与均值差的绝对值， $vp_{downrange}$ 为最小值与均值差的绝对值，

vp_{range} 为最大值与最小值差的绝对值。

基于目标检测的识别结果将各类别异常的位置和置信度转化为类别对应的 30 维空间序列，每个序列点对应 1 米的空间切片，实现了检测框空间分布的有效数字化。具体而言，空间序列特征 $s_c \in R^{30}$ 对应不同类别 c ，通过与施作面完整性特征的融合，构建出表达空间异常强度和语义关联的输入矩阵。

针对空间切片 $j=1,2,\dots,30$ 处的岩体完整性离散等级，本研究构建了多输出随机森林分类模型。模型输出为：

$$\hat{y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_{30}), \hat{y}_j \in \{1,2,3,4\} \quad (7)$$

其中，对每个切片的预测：

$$\hat{y}_j = \arg \max_{c \in \{1,2,3,4\}} P_{j,c}(X) \quad (8)$$

这里 $P_{j,c}(X)$ 为第 j 个切片随机森林分类器预测为类别 c 的概率分布。多输出策略通过训练 30 个独立分类器，实现每米空间切片处的岩体等级分类。

为了更好地体现施作面完整性与其他特征的关联性，模型还引入了交互项，如：

$$wf * V_p^{avg}, wf * vp_{uprange}, wf * clip \quad (9)$$

其中， wf 表示施作面岩体完整性指标， V_p^{avg} 为该区间纵波波速的平均值， $vp_{uprange}$ 为式（6）中定义的纵波波速最大值与平均值之差的绝对值，用于表征纵波波速的上行波动幅度； $clip$ 表示由跨模态模型 CLIP 提取的图像-文本语义匹配相关特征。这些复合特征有效提高了模型对复杂特征间非线性交互的表达能力。

模型训练所用数据集采自实际深部地质工程，格式统一为标准 JSON 文件，包含完整的输入特征和 30 个米段的岩体完整等级标签。本文通过训练过程自动学习不同特征维度的分裂策略和决策边界，相当于在模型内部隐式地完成了对各模态特征有效权重的分配，而非由人工预先指定。

在模型评估阶段，采用分层的 5 折交叉验证，以保持训练与验证数据的代表性与独立性。准确率定义为：

$$Accuracy = \frac{1}{N \times 30} \sum_{i=1}^N \sum_{j=1}^{30} \mathbb{I}(y_{i,j} = \hat{y}_{i,j}) \quad (10)$$

其中 N 为样本总数， $\mathbb{I}(\cdot)$ 是指标函数， $y_{i,j}$ 和 $\hat{y}_{i,j}$ 分别为第 i 个样本第 j 个切片的真实值和预测值。

宏平均 F1 分数是所有类别 F1 分数的算术平均，计算方式如下：

$$F1_{macro} = \frac{1}{C} \sum_{c=1}^C \frac{2Precision_c \cdot Recall_c}{Precision_c + Recall_c} \quad (11)$$

其中 $C=4$ 为类别数, $Precision_c$ 和 $Recall_c$ 为类别 c 的精确率和召回率。

通过该指标, 可以综合衡量模型在不同岩体等级类别上的预测效果, 特别是数据不均衡情况下的分类性能。

模型最终在独立 Hold-out 测试集上验证, 表现稳定且优异, 展示了多源信息融合和多输出随机森林分类器在工程实际中的适用性和推广价值。

在获得沿探测方向岩体完整性预测序列后, 结合前文 2.4 节基于岩石单轴抗压强度、风化程度及岩性分类计算得到的岩石硬度指标, 进一步实现围岩级别的综合判定。具体而言, 将岩体完整性数值与岩石硬度等级作为输入, 通过预先构建的对照表对应映射关系, 获得围岩级别的最终分类结果, 对照规则如表 1 所示。

表 1 实际施工过程围岩级别对照关系表

Table 1 Reference Table of Surrounding Rock Classifications in

Actual Construction Process		
围岩级别	岩石硬度/岩体完整性	存在风险
I级	岩体完整	岩爆
II级	极硬岩: 岩体较完整	岩爆、局部掉块
	硬岩: 岩体完整	局部掉块
III级	极硬岩: 岩体较破碎	掉块
	硬岩: 岩体较完整	掉块、坍塌
	较软岩: 岩体完整	掉块、小坍塌
IV级	极硬岩: 岩体破碎	掉块、小坍塌
	硬岩: 岩体较破碎	掉块、小坍塌
	较软岩: 较完整或较破碎	掉块、坍塌
	软岩: 岩体较完整或完整	掉块、坍塌
V级	较软岩: 岩体破碎	掉块、坍塌、突泥
	软岩: 岩体破碎或较破碎	掉块、坍塌、突泥
	极软岩: 极破碎岩	掉块、坍塌、突泥
VI级	断层; 绿泥石; 软黏土	大型坍塌、突涌水

该对照表依据国家及行业地质工程规范, 结合工程实践经验, 系统整合岩体完整性的连续性指标和岩石硬度的强度属性, 分类为“I级”“II级”“III级”“IV级”及“V级”等多档围岩等级。表中定义了岩体完整性的分段区间与硬度等级的对应规则, 兼顾结构完整性与力学特性的综合表现。

在实际应用中, 针对探测方向沿线各米段的岩体完整性预测值, 首先根据该值所在区间查找对应的围岩完整性等级; 随后利用对应位置的岩石硬度指标进行辅助判定, 二者共同决定最终的围岩级别。

例如, 若某段岩石硬度处于较软等级, 且岩体完整性数值亦较低, 则该段围岩极可能被界定为破碎类, 提示施工风险较高; 相反, 硬度高且完整性良好的区段则归类为良好围岩, 施工条件较优。

3 深地围岩级别预报效果分析

本研究构建的数据集包含 529 组物探图像及其对应的文本响应特征标签, 涵盖约 16280 米的探测范围, 数据来源于 327 次不同的探测过程, 其中 130 组数据同时包含地质雷达探测图像、地震波纵波波速图像及瞬变电磁图像中的 2 至 3 种, 其中 274 组数据仅包含地质雷达探测图像, 其中 57 组数据仅包含瞬变电磁探测图像, 其中 68 组数据仅包含地震波纵波波速图像。通过针对图像及文本、目标框标签的系统数据增强, 最终扩展至 10580 组多模态输入, 有效增加数据样本的多样性。其中, 图像增强采用随机水平翻转、垂直翻转、亮度与对比度调节等操作; 文本增强采用同义词替换及重组搭配的方式; 目标框信息按照图像变化方式同步计算其变化后的位置信息。最终, 训练集和测试集按照 8: 2 的比例划分, 训练过程中模型的收敛趋势及训练效果得到了明显改善。

针对训练完成的模型, 另外选取上述训练集和测试集之外的 23 张图像进行验证测试, 这些验证图像来自 13 次探测, 覆盖约 820 米范围, 其中 18 张图像同时包含地质雷达探测图像、地震波纵波波速图像及瞬变电磁图像中的 2 至 3 种, 2 张图像仅包含地质雷达探测图像, 2 张图像仅包含瞬变电磁探测图像、1 张图像仅包含地震波纵波波速图像。数据经过相同增强处理, 增强后共 460 张物探图像。

为了全面评估所训练的破碎岩体特征检测模型在物探图像中破碎岩体特征检测任务上的性能, 本研究采用了准确率和召回率两个指标。准确率衡量的是预测为正例的目标中有多少是真正的异常目标, 召回率衡量的是所有真实异常目标中被检测出的比例。具体地, 定义如下:

$$precision = \frac{TP}{TP+FP}, recall = \frac{TP}{TP+FN} \tag{12}$$

其中, TP 表示预测框与真实框成功匹配的数量, 匹配条件为类别一致且交并比 (IoU) 达到预设阈值; FP 表示未成功匹配任何真实框的预测框数量; FN 则为未被任何预测框检测到的真实框数量。本研究中, 采用 IoU 阈值 0.5 作为匹配标准, 在测试集的所有数据上统计累加 TP、FP、FN, 最终计算总体准确率

与召回率。在本研究构建的物探图像破碎岩体特征验证集上，模型达到准确率 0.87 和召回率 0.81。破碎岩体特征检测模型对物探图像破碎岩体特征的检测效果如图 4 所示。

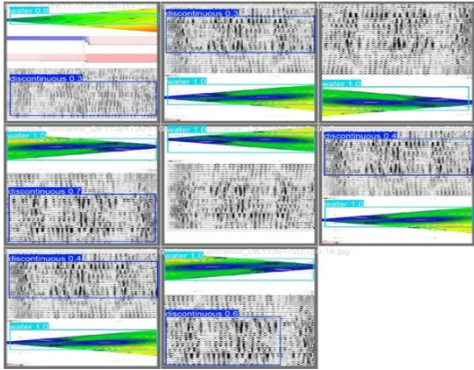


图 4 物探图像破碎岩体特征的检测效果

Fig. 4 Detection Performance of Fractured Rock Mass

需要指出的是，地质雷达图像中“同相轴错断”特征标注依赖专家主观判断，准确性有限，导致该类别识别效果较弱，误报漏报相对较多，准确率约 65%。为弥补这一缺点，本文引入图像-文本跨模态匹配模型，结合专家物探报告文本描述，构建图像-文本对应数据集。通过目标框信息与文本特征共同判断岩体完整性，提高了围岩级别预报的可靠性和准确率。

验证结果表明，针对物探图像的跨模态匹配模型在识别图像响应特征方面表现稳定且准确，整体准确率达到 89%。本文还选取部分典型图像并列出其实际响应特征与模型预测结果的对比情况，详见表 2，可见预测结果与实际标签较为吻合，证明了该模型在物探图像破碎岩体特征跨模态匹配任务中的可靠性和实用价值。

表 2 物探图像跨模态模型响应特征识别效果

Table 2 Recognition Performance of Multimodal Model for Geophysical Image Response Features

物探图像类型	实际标签	预报结论
GPR-TSP-TEM	同相轴不连续-波速	同相轴不连续-波速
	下降-较高阻	下降-较高阻
GPR	同相轴不连续	同相轴连续性差
GPR-TSP	同相轴连续性差-波	同相轴连续性差-波
	速下降	速下降
GPR-TEM	同相轴连续性一般-	同相轴连续性一般-
	较高阻	较高阻
TSP	波速下降	波速下降
TEM	较低阻	较低阻

本次选取的测试段落来自几个不同的深部地质工程施工区域，包含多种地质情况，覆盖Ⅲ级、Ⅳ级、Ⅴ级围岩。融合破碎岩体响应特征文本信息、破碎岩体特征目标框信息、纵波波速最大值 V_p^{max} 、纵波波速均值 V_p^{avg} 、纵波波速最小值 V_p^{min} 、施作面地质情况、岩石岩性及岩石风化程度多模态地质信息进行围岩级别预报，在测试段内本文所提方法的围岩级别预报准确率达到 92%，其中，Ⅲ级围岩段落预报准确率为 95%、Ⅳ级围岩段落预报准确率为 86%、Ⅴ级围岩段落预报准确率为 88%。以测试集中某两张物探图像的探测信息为例，深部地质工程中实际开挖揭示的围岩级别分布情况、各方法预报得到的围岩级别情况如图 5 所示，该图展示了第一个探测区域和第二个探测区域内实际开挖揭示的围岩级别分布情况、文本所提方法的围岩级别预报情况、常用方法的围岩级别预报情况及消融实验中围岩级别预报情况。本文在统计围岩级别预报准确率时均以米为单位，计算预报正确里程占总里程的百分比。当前模型的泛化能力主要在Ⅲ级别、Ⅳ级、Ⅴ级围岩地质条件下得到了验证，对其他类型地质构造区，如大型断层、强岩溶发育区等的适用性仍有待在多工程、多区段数据集上进一步检验。

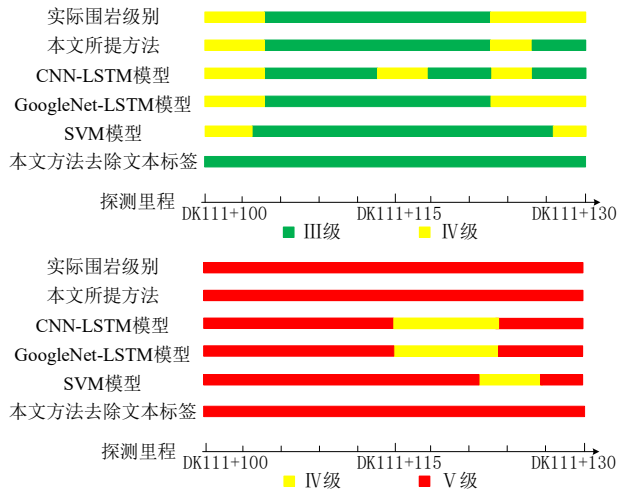


图 5 某两段探测里程内围岩级别实际分布情况及预报效果

Fig. 5 Actual Distribution and Prediction Performance of Surrounding Rock Grades within Two Survey Sections

本文将围岩级别视作沿里程方向的一维离散场，采用 LSTM 对相邻里程单元间的空间连续性进行建模，故选用 CNN-LSTM 与 GoogLeNet-LSTM 作为针对序列输出任务的强基线模型。基于这两种常用方法，同样是融合破碎岩体响应特征文本信息、破碎岩体特征目标框信息、 V_p^{max} 、 V_p^{avg} 、 V_p^{min} 、施作面地质情况、岩石岩性及岩石风化程度多模态地质

信息进行围岩级别预报，CNN-LSTM 模型在 23 张物探图像中的预报准确率为 84%；GoogleNet-LSTM 模型在 23 张物探图像中的预报准确率为 87%。在处理混合类型变量和结构化变量时，两者容易出现过拟合现象，而本文所提方法在保持较高准确率的同时表现出更强的鲁棒性。此外，本文还构建了基于 SVM 模型的对比实验，将文本特征、目标框信息、 V_p^{max} 、 V_p^{avg} 、 V_p^{min} 、施作面地质情况、岩石岩性及岩石风化程度统一编码为固定维度特征向量作为模型输入。为适配 SVM 的单输出特性，本文将 30m 范围内每 1m 的围岩级别预测拆分为 30 个并行的分类子任务，对每个里程位置分别训练一个 SVM 分类器。最终，得到的预报准确率为 79%。本研究的输入特征是多模态的，包括文本特征、图像特征和地质数值特征，SVM 在处理这类异构数据时可能无法充分捕捉它们之间的复杂关系。各方法的对比结果见表 3。

表 3 实际隧道段落内围岩级别预报效果

Table 3 Prediction Performance of Surrounding Rock Grades

within Actual Tunnel Sections	
预报方法	预报准确率
本文所提方法	92%
CNN-LSTM 模型	84%
GoogleNet-LSTM 模型	87%
SVM 模型	79%
本文所提方法去除文本标签	89%

当前深部地质工程领域常用的针对物探图像的围岩级别预报方法通常未使用文本标签，因此本实验去除文本标签，仅融合破碎岩体特征目标框信息、 V_p^{max} 、 V_p^{avg} 、 V_p^{min} 、施作面地质情况、岩石岩性及岩石风化程度进行围岩级别预报，采用本文使用的随机森林分类模型，在相同探测里程内进行围岩级别预报。去除文本标签信息后，模型在 23 张物探图像中的预报准确率为 89%。当地质雷达图像中出现同相轴错断或同相轴不连续等特征时，模型容易误判或漏报。例如，当同相轴不连续出现在图像中时，模型往往将其漏报为同相轴连续性差，而当出现同相轴连续性差时，模型可能会误报为同相轴不连续。这种情况导致了模型的较高误报和漏报率。然而，经过图文跨模态匹配模型与目标检测模型的互相修正，模型的准确性显著提高。单独使用目标检测模型时，漏报率为 30%，误报率为 40%；而在引入文本信息后，漏报率降低至 15%，误报率下降至 25%，证明了文本标签在减少误报和漏报方面的有效性。

可见，去除文本标签后，模型的预报准确率有明显下降，文本标签可以和目标框标签相互补充、修正，从而提高围岩级别预报准确率。

4 结论

针对深部地质工程围岩级别判定过程中存在的物探图像解译主观性强、破碎岩体特征难量化及单一模态信息利用不充分等问题，本文提出了一种融合图像-文本多模态信息的围岩级别智能建模及预报方法。该方法通过挖掘物探报告中蕴含的专家先验知识，构建了图像响应特征与文本语义描述的映射关系，并结合多源物探数据的物理参数，实现了对围岩级别的智能化、高精度判定。主要贡献体现在：1) 本文提出一种融合图像-文本多模态探测信息的围岩级别智能预报方法，能够实现多源物探图像到探测里程范围内的围岩级别预报，预报结论可靠、可解释性强；2) 针对多源物探图像破碎岩体响应形式差异大、特征表现复杂的挑战，构建基于目标检测与跨模态语义约束的异常体识别方法，有效解决工程语义难以量化的问题。

经验证，在覆盖约 820m 探测里程、包含多种地质工况的测试集上，本文方法的围岩级别预报准确率达到 92%。对比实验结果表明，该方法显著优于主流的时空序列预测模型，相比于 CNN-LSTM 模型提升了 8%，相比于 GoogleNet-LSTM 模型提升了 5%。这证明了本文所设计的特征提取与定位模型在处理复杂地质信息时具有更强的鲁棒性与特征表达能力。消融实验结果显示，在去除文本标签信息仅保留视觉特征与物理参数的情况下，模型的预报准确率下降至 89%。这表明引入物探报告中的专家文本语义信息能够有效修正视觉模型在模糊特征下的误判，带来了 3%的精度提升。文本信息作为一种强先验约束，有效弥补了仅靠图像像素特征识别破碎岩体时的多解性缺陷。

综上所述，本文提出的融合图像-文本多模态探测信息的围岩级别智能预报方法，不仅实现了从定性解译向定量预报的跨越，更为深部超前地质预报提供了一种兼顾物理机理与专家经验的智能化解决方案，具有重要的工程应用价值。

References

Zou Y., Dong X. J., Feng T., et al., 2025. Research on spatial prediction technology for mitigating tunnel inrush disasters under complex geological conditions in China's Hengduan Mountain Range, *Scientific Reports*, 15(1): 1850-1850.

- Sun H., Sheng L., Dai Y., et al., 2025. 3D geological modeling of tunnel alignment in the complex mountainous region of Yongshan, China, based on multisource data fusion, *Engineering Geology*, 354: 108209-108209.
- Sabri S. M., Jaiswal A., Verma K. A., et al., 2025. Systematic Review of RMR, Q-System, and GSI in Tunnel Classification: Origin, Advancement, and Limitations, *Indian Geotechnical Journal*, (prepublish): 1-23.
- Li X., Zhang X., Liu Q., et al., 2025. Evaluation of rock mass quality and its mechanical properties through digital drilling process monitoring, *Journal of Rock Mechanics and Geotechnical Engineering*, 17(7): 4490-4511.
- Zhang Y., Zhou J., Li J., et al., 2025. Advancing overbreak prediction in drilling and blasting tunnel using MVO, SSA and HHO-based SVM models with interpretability analysis, *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, 11(1): 53-53.
- Gan C., Wang Y., Cao W. H., et al., 2025. Real-time formation drillability sensing-based hybrid online prediction method for the rate of penetration (ROP) and its industrial application for drilling processes, *Control Engineering Practice*, 164: 106487-106487.
- Zhao Z. P., Chen J. X., Geng Q., et al., 2024. Construction Technology and Application of TBM in Shengli Tunnel of Tianshan Mountain under poor geological conditions, *Stavební Obzor Civil Engineering Journal*, 33(2): 168-180.
- Liu C., Li J., Liu Z., et al., 2025. A comprehensive review of data processing and target recognition methods for ground penetrating radar underground pipeline B-scan data, *Discover Applied Sciences*, 7(4): 310-310.
- Zheng P., Zhang A., Shi Z., et al., 2025. TLAD-YOLO: Lightweight network for intelligent detection of railway tunnel lining anomalies using ground penetrating radar, *Journal of Applied Geophysics*, 241: 105869-105869.
- Dong Z., Zhang X., Yang W., et al., 2025. Using a Hybrid of the U-Net and Multilayer Perceptron Models to Automatically Predict the Geological Strength Index from Tunnel Face Rock Joint Images, *Rock Mechanics and Rock Engineering*, 58(11): 1-18.
- Yu X. Y., Dong N., Zhu L. H., et al., 2025. CLIP-Driven Semantic Discovery Network for Visible-Infrared Person Re-Identification, *IEEE Transactions on Multimedia*, 27: 4137-4150.
- An J., Su P., Liu J., et al., 2025. Implementation of YOLO-CLIP fusion algorithm for fall detection, *Signal, Image and Video Processing*, 19(10): 837-837.
- Saeed F., Aldera S., Shamma'a A. A. A., et al., 2024. Rapid adaptation in photovoltaic defect detection: Integrating CLIP with YOLOv8n for efficient learning, *Energy Reports*, 12: 5383-5395.
- Li J. F., Sun S. Y., Zhang K., et al., 2025. Single-stage zero-shot object detection network based on CLIP And pseudo-labeling, *International Journal of Machine Learning and Cybernetics*, 16(2): 1055-1070.
- Xu Z., Wang Z., Li S., et al., 2024. GeoPredict-LLM: Intelligent tunnel advanced geological prediction by reprogramming large language models, *Intelligent Geoengineering*, 1(1): 49-57.
- Fang Z. H., Tan Z. H., Zhu Z., et al., 2023. Detection of highway disease images using ground-penetrating radar based on Faster R-CNN. *Construction Technology (Chinese and English)*, 52(24): 76-82. (in Chinese)
- Qin Z. X., Jiang Y. N., Xu L., et al., 2023. Automatic anomaly detection of ground-penetrating radar road images based on YOLO algorithm. *Science Technology and Engineering*, 23(27): 11505-11512. (in Chinese)
- Wang W. W., Mou Y. C., 2021. Application of TSP tunnel advanced geological forecasting technology in large faults. *Journal of Engineering Geophysics*, 18(2): 186-193. (in Chinese)
- Zhang X. J., 2023. Application of transient electromagnetic method in advanced geological forecasting of water diversion tunnels. *Heilongjiang Traffic Science and Technology*, 46(10): 102-105. (in Chinese)
- Zhang S., Zhen D. Y., Liu K., et al., 2024. Application of 3D visualization of TSP and transient electromagnetic results in tunnel fault water-bearing forecasting. *China Railway*, (5): 75-81. (in Chinese)
- Lin Z., Feng S., Fan G. Y., et al., 2021. Application of geological radar in advanced geological forecasting of a tunnel. *Exploration Science and Technology*, (3): 47-51+64. (in Chinese)
- Gao S. Q., Wang Y. Q., Mou Y. C., 2021. Advanced geological forecasting and image analysis of karst tunnels using geological radar. *Journal of Engineering Geophysics*, 18(5): 642-646. (in Chinese)

- Long G. Y., Zheng L., 2024. Research on surrounding rock classification based on tunnel advanced geological detection. *Science and Technology Innovation and Application*, 14(6): 124-127. (in Chinese)
- Jiang Y., Wang H. L., Chen Z., 2024. Intelligent prediction algorithm for advanced forecasting images of adverse tunnel geological bodies based on deep learning. *Modern Tunnel Technology*, 61(3): 148-156. (in Chinese)

中文参考文献

- 房振华, 谭治海, 朱哲, 等, 2023. 基于 Faster-RCNN 的探地雷达公路病害图像检测, *施工技术(中英文)*, 52(24): 76-82.
- 覃紫馨, 姜彦南, 徐立, 等, 2023. 基于 YOLO 算法的探地雷达道路图像异常自动检测, *科学技术与工程*, 23(27): 11505-11512.
- 王汪汪, 牟元存, 2021. TSP 隧道超前地质预报技术在宽大断层内部的应用, *工程地球物理学报*, 18(2): 186-193.
- 张旭杰, 2023. 瞬变电磁法在引水隧洞超前地质预报中的应用, *黑龙江交通科技*, 46(10): 102-105.
- 张硕, 甄大勇, 刘康, 等, 2024. TSP 与瞬变电磁成果三维化在隧道断层含水预报中的应用, *中国铁路*, (5): 75-81.
- 林志, 冯森, 范国宇, 等, 2021. 地质雷达在某隧道超前地质预报中的应用, *勘察科学技术*, (3): 47-51+64.
- 高树全, 王玉琴, 牟元存, 2021. 岩溶隧道地质雷达超前地质预报及图像分析, *工程地球物理学报*, 18(5): 642-646.
- 龙贵云, 郑亮, 2024. 基于隧道超前地质探测的围岩分级研究, *科技创新与应用*, 14(6): 124-127.
- 蒋源, 王海林, 陈兆, 2024. 基于深度学习的隧道不良地质体超前预报图像智能预测算法, *现代隧道技术*, 61(3): 148-156.