

doi:10.3963/j.issn.1001-487X.2025.04.002

## 基于 MDO-XGBoost 算法的钻凿参数与 岩体等级关系研究\*

曹贞洋<sup>1</sup>, 贾家银<sup>2</sup>, 龚敏<sup>1</sup>, 赵新祥<sup>2</sup>, 吴昊骏<sup>1</sup>, 周世均<sup>2</sup>, 艾卿<sup>2</sup>, 高兴<sup>2</sup>

(1. 北京科技大学资源与安全工程学院, 北京 100083; 2. 重庆中环建设有限公司, 重庆 401120)

**摘要:** 根据岩体实际等级优化爆破方案是提高爆破效果的一个重要途径,但在现场施工时难以直接获取开挖面各部分岩体的等级信息。为此以重庆某隧道为背景,提出一种基于现场钻凿参数识别岩体等级的方法。现场勘测岩体特征并划分岩体等级,采集钻凿数据并筛选出与岩体等级相关的钻凿参数;针对不同岩体等级的样本数量不均衡情况,采用基于马氏距离的过采样技术(MDO)扩充小样本类别的数据量;利用极端梯度提升(XGBoost)算法拟合钻凿参数与岩体等级的关系,建立岩体等级识别模型;最后基于模型识别结果优化现场爆破方案。结果表明:MDO-XGBoost模型对岩体等级整体识别率达到80%;基于岩体等级识别结果优化的爆破方案提高了爆破进尺,且缩短了出渣时间。研究提供了一种在样本不均衡时基于钻凿参数准确识别岩体等级的方法,有助于隧道现场的智能爆破和高效施工。

**关键词:** 钻凿参数; 岩体等级识别模型; 不均衡样本; 过采样技术; 极端梯度提升算法

**中图分类号:** U456.3 **文献标识码:** A **文章编号:** 1001-487X(2025)04-0009-13

## Relationship between Drilling Parameters and Rock Mass Classification based on MDO-XGBoost Algorithm

CAO Zhen-yang<sup>1</sup>, JIA jia-yin<sup>2</sup>, GONG Min<sup>1</sup>, ZHAO Xin-xiang<sup>2</sup>,  
WU Hao-jun<sup>1</sup>, ZHOU Shi-jun<sup>2</sup>, AI Qing<sup>2</sup>, GAO Xing<sup>2</sup>

(1. School of Resources and Safety Engineering, University of Science and Technology Beijing, Beijing 100083, China; 2. Chongqing Zhonghuan Construction Co., Ltd., Chongqing 401120, China)

**Abstract:** Optimizing blasting plans according to actual rock mass classification represents a critical approach for enhancing blasting outcomes. However, field construction conditions present significant challenges in obtaining direct rock mass classification data across different excavation face zones. This study proposes a rock mass classification method utilizing on-site drilling parameters, with application to a Chongqing tunnel project. Firstly, the characteristics of the rock mass were investigated, and the rock mass class was divided. The drilling data were then collected, and drilling parameters related to the rock mass class were screened. Furthermore, the data volume of small sample classes was expanded by the Mahalanobis Distance-based Over-sampling technique (MDO). Meanwhile, the relationship between drilling parameters and rock mass class was modeled using the Extreme Gradient Boosting (XGBoost) algorithm, and an identification model for rock mass class was then established. Finally, the site blasting scheme was opti-

收稿日期 (Date of reception): 2025-02-20

网络首发日期 (Published online): 2025-05-15

作者简介: 曹贞洋(1996-), 男, 安徽金寨人, 博士研究生在读, 从事岩石动力学、工程爆破方面的研究, (E-mail) D202110001@xs.ustb.edu.cn。

通信作者: 龚敏(1963-), 男, 重庆巫山人, 教授、博士研究生导师, 从事岩石动力学、工程爆破方面的研究, (E-mail) gongmustb@163.com。

基金项目: 北京市科学技术协会青年人才托举工程 (No. BYESS2023043)

**About the author:** CAO Zhen-yang (1996-), male, born in Jinzhai city, Anhui province, doctoral candidate, mainly engaged in rock dynamics and engineering blasting, (E-mail) D202110001@xs.ustb.edu.cn.

**Corresponding author:** GONG Min (1963-), male, born in Wushan city, Chongqing municipality, professor, doctoral supervisor, mainly engaged in rock dynamics and engineering blasting, (E-mail) gongmustb@163.com.

**Fund Programs:** the Young Elite Scientists Sponsorship Program by BAST (No. BYESS2023043)

mized based on the identification results. The results show that the MDO-XGBoost model achieves an overall classification accuracy of 80% for rock mass grade identification. The optimized blasting scheme has increased blasting penetration and shortened deslagging time, based on the rock mass classification results. This research presents a viable method for accurately identifying rock mass classes based on drilling parameters, particularly in the context of sample imbalance, thereby contributing to intelligent blasting and efficient construction in tunnel projects.

**Key words:** drilling parameters; identification model of rock mass classification; imbalanced samples; over-sampling technology; extreme gradient boosting

钻爆法开挖广泛应用于隧道、竖井、巷道等地下工程建设中,其中钻孔作业中钻机破坏岩体的过程在力学上可认为是钻头与岩体的相互作用,各项钻凿数据与岩体质量存在着一定的对应关系<sup>[1,2]</sup>。明确这一关系并建立基于钻凿参数的岩体等级识别模型,有助于在开挖过程中感知前方岩体特征<sup>[3]</sup>、调整爆破方案<sup>[4]</sup>、优化施工工序<sup>[5]</sup>、保障安全高效生产<sup>[6]</sup>。基于钻凿参数识别岩体等级、进而实时调整爆破参数,是智能爆破建设的核心和重要组成部分。目前,能够实时采集钻凿参数的电脑凿岩台车在工程上逐渐得到广泛的应用,为研究钻凿参数与岩体等级关系提供了有力的技术支撑。

现场钻凿参数与岩体等级的关系属于多元非线性关系<sup>[7]</sup>,采用以机器学习为代表的人工智能算法建立这类关系模型具有技术优势。易文豪等采集了 299 份钻孔样本,利用支持向量机(SVM)建立了掌子面围岩分级模型,在样本数量较少条件下达到较好的准确度<sup>[8]</sup>。张幼振等采用了 C 均值聚类算法训练钻进速度、转速和钻进压力等钻凿参数<sup>[9]</sup>,实现了含煤地层的岩性识别。Rajib Ghosh 利用主成分分析建立了钻进数据与岩体裂隙、空腔之间的关系模型<sup>[10]</sup>,并通过模型识别结果来指导后续装药。王明年等采用极端随机树算法训练 II ~ IV 级围岩的钻凿数据<sup>[11]</sup>,建立智能分级模型,促进了在三维空间上围岩等级的划分。其他学者也利用智能算法开展了钻凿参数或掘进参数与岩体特征的研究<sup>[12-14]</sup>。

在基于现场试验研究钻凿参数-岩体等级关系时,由于地质勘探的条件有限,难以在不同岩体环境下采集等量的大规模样本,经常存在着各类别样本不均衡的问题。这种数据的不均衡性将极大地影响岩性识别模型在少样本类别的识别率<sup>[15]</sup>。过采样技术是当前处理样本不均衡情况的一种方法,LIU 等应用合成少数类的过采样技术(SMOTE)解决数据库中岩体分类比例不平衡的问题<sup>[16]</sup>,进而提高模型在少数类上的识别率;王光宇等在构建随机森林分类器前为提高岩性数据与地震资料的吻合度<sup>[17]</sup>,利用 NM-SMOTE 算法来平衡岩性样本集。但值得

注意的是,在利用过采样算法合成新样本时,往往会由于未考虑到各类样本的分布区域,造成新合成的少数类样本扩展到了多数类样本群中,增大了不同样本之间重叠的概率<sup>[18]</sup>。机器学习中的集成学习算法也可以提高模型在不均衡样本集中的泛化能力<sup>[19,20]</sup>,在地应力反演<sup>[21]</sup>、边坡稳定性预测<sup>[22]</sup>、矿区岩层探测等领域均有应用<sup>[23]</sup>。但是单独采用某一种方法,对样本不均衡时的识别效果提升仍是有限的。

为构建现场钻凿参数与岩体等级的关系模型,提出联合采用基于马氏距离的过采样技术(MDO)和集成学习模型 XGboost 两种算法<sup>[24,25]</sup>,来解决样本不均衡导致的识别准确率不高的问题。在获取到现场样本不均衡的数据集后,首先采用 MDO 算法处理样本分类不均衡的问题,在扩充少数类样本量的同时,维持了样本的分布特征,避免过采样造成的不同类样本的重叠;然后利用 XGBoost 算法分析过采样后各等级岩体的钻凿数据,通过集成学习的方法增强模型的整体表现,进一步削弱样本不均衡对模型性能的影响,构建基于钻凿参数识别岩体等级的模型。在样本不均衡情况下采用上述方法,可提高所构建模型对岩体等级识别的准确率,有助于现场准确识别岩体特征、及时调整爆破参数、优化爆破方案。

## 1 研究路线与方法

### 1.1 研究路线

获取现场不同岩体等级条件下的钻凿参数数据,是研究钻凿参数与岩体等级的关系、建立岩体等级识别模型的基础。其中,钻凿参数数据可利用电脑凿岩台车传感系统采集得到,岩体等级信息需要在现场开展地质勘探获取。但是,对于某些等级(如 I 级、V 级)岩体,由于现场条件的限制,在开挖区段能够采集到的样本量较少,导致各岩体等级的样本在数量上不均衡。为了在这种不均衡条件下,通过钻凿参数的变化较准确地识别岩体等级,提出利用 MDO 过采样算法处理现场各等级样本不均衡

的问题,并利用机器学习算法 XGBoost 训练过采样后的样本,构建岩体等级识别模型。图 1 为本文的研究路线,主要包括数据采集、特征筛选、过采样处理和模型构建 4 个部分。

首先在试验区段勘探地质情况,确定各区段岩体的主要特性,根据《工程岩体分级》(GB/T 50218—2014),计算岩体基本质量指标(BQ 值);同

时,在现场不同地质区段上开展钻凿试验,通过电脑凿岩台车的传感系统获取凿岩台车工作数据。为了减少冗余数据和不相关参数,计算各种钻凿参数和岩体等级的相关性系数,并分析不同岩体等级条件下各钻凿参数的数据分布情况,筛选出应用于后续模型构建的主要钻凿参数,形成不同等级岩体条件下的钻凿参数数据集。

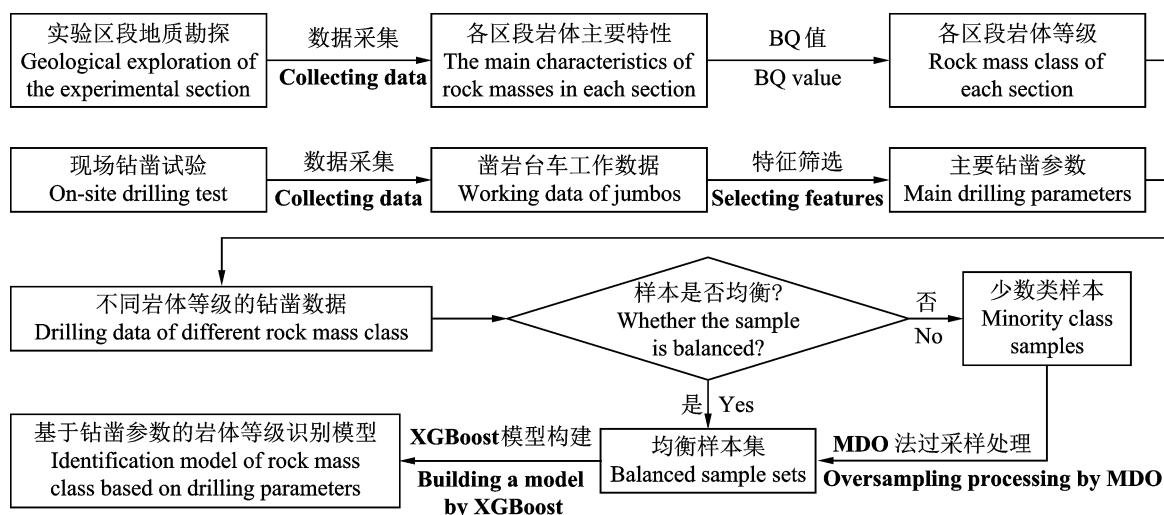


图 1 研究路线

Fig. 1 Research route

面对不同岩体等级的钻凿数据,需要判断其样本数量是否均衡,对于各类别数量均衡的样本,将其直接用于后续的模式构建;对于各类别的数量不均衡的样本,从中提取出数量较少的少数类样本,利用 MDO 法进行过采样处理,合成足够量的少数类样本,形成各类别数量均衡的样本集。当样本集中各等级岩体的样本量均衡时,以主要钻凿参数为输入特征、岩体等级为输出值,采用 XGBoost 算法拟合各主要钻凿参数与岩体等级的关系,构建岩体等级识别模型。

在上述四个环节中,MDO 过采样算法与 XGBoost 算法是本文处理现场样本不均衡问题、构建岩体等级识别模型所应用的重要方法,以下对两种算法的基本原理作具体介绍。

### 1.2 MDO 过采样数据处理方法

MDO(Mahalanobis Distance-based Over-sampling technique)算法是 Lida ABDI 提出的一种考虑到样本数据在各维度上分布特征的过采样技术<sup>[26]</sup>,用于处理数据分析过程中各类样本数量不均衡的问题。该算法首先根据实测样本分布的概率轮廓,提取合适的样本进行后续过采样操作,减少了不同类别的样本区域重叠的风险;在对少数类样本进行过采样

时,采用马氏距离来度量样本与样本均值的距离,保留了少数类实测样本的协方差结构,从而降低过采样操作对各类样本数据结构的影响。图 2 为 MDO 算法过采样的流程图,主要包括以下 4 个步骤:合适样本的提取、样本空间的转换、样本过采样、样本空间的恢复。在完成上述 4 个步骤后,少数类样本完成了过采样操作,数据可用于后续的模式训练。

### 1.3 XGBoost 算法原理

XGBoost 是一种在 GBDT 算法基础上改进的集成算法,将多个弱学习器(决策树)整合为一个强学习器<sup>[27]</sup>。在每次迭代训练一棵新的决策树时,XGBoost 通过计算损失函数梯度,来最小化当前模型在训练数据上的残差,逐步改善模型整体的预测效果;为提高模型泛化能力,XGBoost 算法引入了  $L_1$  和  $L_2$  正则项,并提供早停机制,以防止过拟合。同时,XGBoost 参照随机森林的做法,每次在节点分裂前进行了随机采样,增加了模型多样性。

## 2 现场数据的采集与特征筛选

### 2.1 工程背景

研究依托于重庆高速集团大巴山隧道工程,隧道总长 13.6 km,其中重庆境内段 5.4 km,最大埋深

达到 1114 m,隧道通过断层破碎带、岩溶等多种复杂地质,包括灰岩、白云岩等多种岩性。主要采用钻

爆法开挖,钻孔工作由 ZYS113 电脑凿岩台车(如图 3)完成。

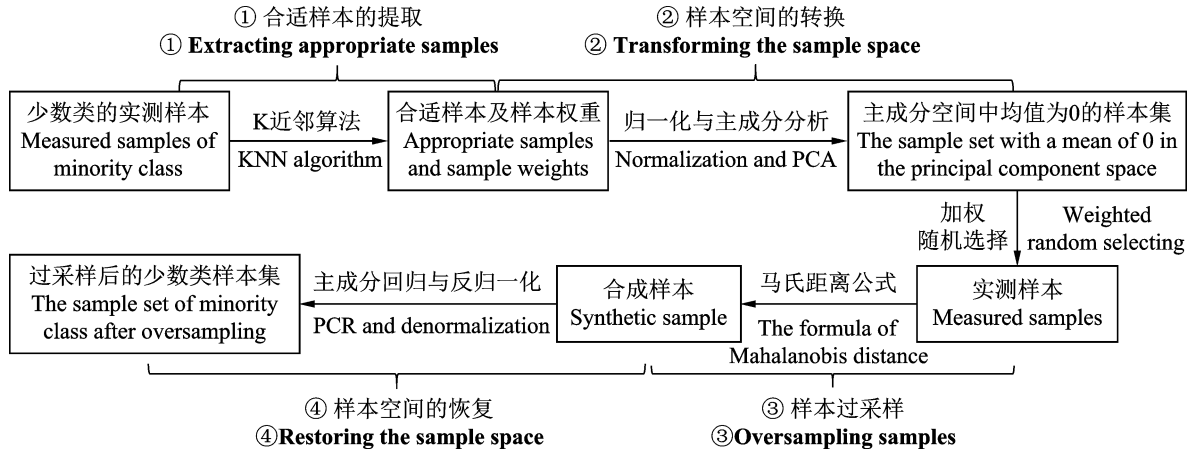


图 2 MDO 算法过采样流程图

Fig. 2 Flow chart of MDO algorithm oversampling

若在隧道开挖过程中,能够及时识别开挖面的岩性变化,并根据岩体变化来调整施工方案,将有效保障安全生产、保证施工效率。利用地质勘探获取开挖面岩体特征是工程上的常用方法,但考虑到经济成本和勘探效率,很难在特长隧道的全路段上开展。因此,提出基于现场电脑凿岩台车的钻凿数据来识别岩体等级的方法。

岩芯采样点均在同一断面上,且距离不超过 2 m,使所采集到的岩体特征数据与钻凿数据来源于同一岩体(如图 4)。



图 3 现场作业的电脑凿岩台车

Fig. 3 A working intelligent drill jambo

## 2.2 现场岩体特征勘测与钻凿数据采集

为分析现场样本数据、研究钻凿参数与岩体等级的关系,并构建岩体等级识别模型,现场需要采集钻凿数据和岩体特征数据,其中存在以下两个方面问题需要解决:钻凿数据可通过电脑凿岩台车自带的传感系统按照 0.02 m 的采样间隔采集到大量样本,但工作参数的种类较多,需对其进行筛选;另一方面,由于现场条件限制,岩体特征数据无法随时随地获取,故选取部分区段勘测岩体特征。

本次研究在不同岩性的区段开展多次钻凿试验与取芯试验。为确保所获得的凿岩参数和后续测定的岩体等级相互对应,每次试验中的钻凿试验点和

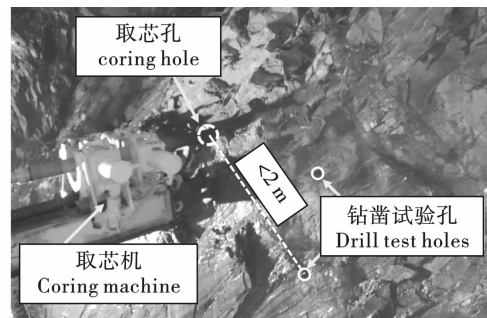


图 4 现场取芯孔与钻凿试验孔的位置关系

Fig. 4 The position relationship between the field core hole and the drilling test holes

为准确获取试验区段的岩体特征和岩体等级,在现场试验过程中采用多功能钻机钻取岩芯(如图 5),利用所取岩芯在实验室开展力学实验(如图 6),测定其岩石强度;利用超声波检测的手段测定钻凿试验孔处岩体的完整程度(如图 7)。按照《工程岩体分级标准》中的岩体基本质量分级方法(BQ法)将试验区段划分为 I ~ IV 级岩体<sup>[28]</sup>,表 1 展示了现场试验测得的岩石强度、岩体完整性系数和岩石基本质量指标的区间以及计算所得各等级岩体 BQ 值区间。

现场采集凿岩台车在上述四类岩体条件下的有效工作数据共 12 203 组,根据每次试验所测定的岩体等级对钻凿数据进行分类,得到 I 级围岩下的钻凿数据 1088 组,II 级围岩 4181 组,III 级围岩的数据

5651 组,IV级围岩的数据 1276 组,四类岩体下的数据量占比约为 1:3.84:5.19:1.17,各类样本之间存在不均衡情况。相比于 III 级围岩充足的数据量, I 级和 IV 级围岩属于小样本类别,后续建立模型时需充分考虑小样本类别对模型识别性能的影响。



图 5 现场钻取的部分岩芯  
Fig. 5 The core drilled at the site



图 6 实验室单轴压缩实验  
Fig. 6 Laboratory uniaxial compression experiment

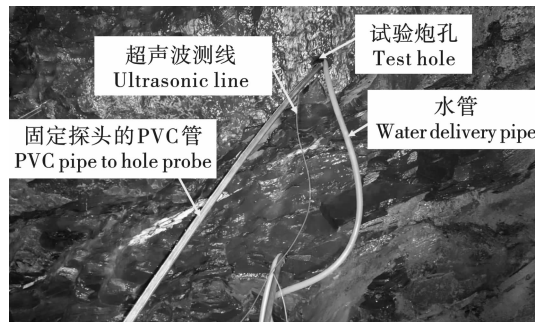


图 7 超声波检测  
Fig. 7 Ultrasonic detection

### 2.3 钻凿数据分析与特征参数筛选

现场采集所得电脑凿岩台车钻凿数据中的工作数据包括钻进速度 ( $v$ )、冲击压强 ( $P_h$ )、推进压强 ( $P_f$ )、旋转压强 ( $P_r$ )、推进速度阀门开度 ( $V_v$ )、旋转速度阀门开度 ( $V_r$ )、推进压强阀门开度 ( $V_f$ )、冲击压强阀门开度 ( $V_h$ ) 等共 8 种参数。各参数的变化与现场岩体条件有关,同时钻凿参数之间具有一定的相关性。为降低参数冗余,更加清晰地反映岩体特征与钻凿参数之间的关系,采用相关性分析的方法,通过 Spearman 相关性系数来评价各钻凿参数与岩体等级以及钻凿参数之间的相关性,根据相关性程度筛选冗余参数。

表 1 隧道试验区段岩体特征及等级

Table 1 Rock mass characteristics and category in tunnelling test sections

序号 Serial number	单轴抗压强度 UCS/MPa	完整性系数 $K_v$	岩石基本质量 指标 $RQD/\%$	岩体基本 质量指标 $BQ$	岩体等级 Rock mass class
1	[132,159]	[0.76,0.79]	[76.68,78.24]	[554,600]	I
2	[120,131]	[0.57,0.60]	[40.61,56.97]	[486,502]	II
3	[60,132]	[0.35,0.57]	[10.56,46.00]	[372,466]	III
4	[47,90]	[0.15,0.20]	[0,5.65]	[268,294]	IV

图 8 为各参数之间的相关性系数情况,与岩体等级相关度最高的钻凿参数为推进压强;在不同等级岩体中开展作业时,为保证钻孔作业能够顺利开展,操作人员会不断调节其他参数而促使钻进速度保持在一个相对稳定的范围内。虽然钻进速度作为钻头与岩石作用中的重要指标参数,但与各参数之间的关系并非简单的线性关系,故其与各参数的 Spearman 相关性系数均不高。

推进压强与推进速度阀门开度、推进压强阀门开度的相关性系数均大于 0.8,冲击压强与冲击压强阀门开度的相关性也达到了 0.84。相关性高的钻凿参数在数据特征上一致性较高,存在着信息冗余。因此,在保留推进压强、冲击压强的前提下,筛除推进

速度阀门开度、推进压强阀门开度和冲击压强阀门开度。下面主要分析钻进速度、冲击压强、推进压强、旋转速度阀门开度、旋转压强等 5 项参数数据分布情况。

图 9 为各项钻凿参数在不同岩体条件下的数据分布情况。由图 9(a) 可得,各等级岩体条件下的大部分数据的钻进速度均集中在 2.5 ~ 5 m/min,这表明现场作业将钻进速度维持在了一个适中范围内,避免速度过慢影响效率或速度过快造成卡钻;由图 9(b) 可得,冲击压强均集中在 140 ~ 160 bar 的区间范围内,维持在较高值域以持续提供破岩动力。

旋转速度阀门开度在一定程度上代表着钻头转速的大小,由图 9(c) 可知转速阀门开度在 I 级岩体条件较其他等级岩体更加集中,且转速减慢。由于

Ⅲ、Ⅳ级岩体完整性较差,岩体中的裂隙、空洞导致转速发生较大变化,在Ⅲ、Ⅳ级岩体中转速阀门开度的波动性更大。

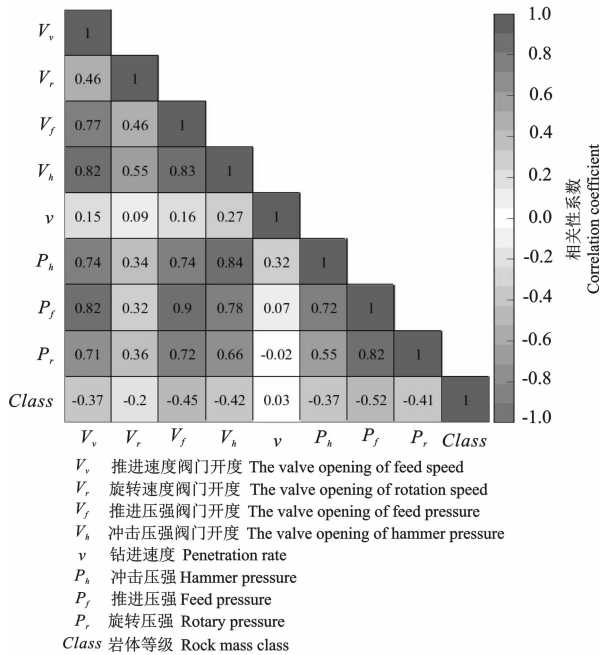


图8 参数相关系数矩阵

Fig. 8 Correlation coefficient matrix of parameters

由图9(d)和图9(e)可知,旋转压强和推进压强的数据分布在不同等级岩体中有较为明显的变化:两类压强在Ⅲ、Ⅳ级岩体中主要分布在较低值域,在Ⅰ、Ⅱ级岩体中取值较高。这一分布趋势在推进压强中最为明显:当岩体等级为Ⅰ级时,60%样本数据的推进压强分布在80~90 bar的高值范围内;而当岩体等级为Ⅳ级时,86%样本的推进压强取值在50 bar以下。Ⅰ、Ⅱ级岩体较Ⅲ、Ⅳ级岩体强度更高、完整性更好,故破碎岩石所需要扭矩和推进力更大。

综上可知,钻进速度、冲击压强、旋转速度(以旋转速度的阀门开度表示)、旋转压强、推进压强均与岩体等级存在不同程度的关系,不同岩体条件下的钻凿参数的数据分布不同。可结合上述五种钻凿参数分布特征,采用人工智能算法拟合钻凿参数与岩体等级之间的关系,构建岩体等级识别模型。

### 3 岩体等级识别模型构建与评估

由于钻凿参数与岩体特征之间的复杂关系难以用单个函数关系式表述,故采用人工智能算法构建岩体等级识别模型,以建立钻凿参数与岩体等级之间的关系模型,进而实现对现场岩体等级的感知。

#### 3.1 数据集划分及样本不均衡问题的处理

构建岩体等级识别模型首先按照一定比例将各类岩体的数据划分为训练集和测试集:训练集数据用于拟合钻凿参数与岩体等级之间的关系、构建并优化模型;测试集数据不参与模型的构建和训练过程,只用于评估最终所得模型对各类岩体的识别性能。本文采用了模型训练中常用的8:2比例。表2为训练集和测试集划分后的数据量情况。

表2 训练集与测试集的数据分配

Table 2 Data allocation of training set and test set

岩体等级 Rock mass class	数据总量 Data volume	训练集 Training set	测试集 Test set
Ⅰ	1088	864	224
Ⅱ	4181	3346	825
Ⅲ	5651	4524	1127
Ⅳ	1276	1022	254
合计 Totality	12203	9761	2442

根据表2可知,各类岩体的样本量存在不均衡情况,Ⅰ级岩体和Ⅳ级岩体的数据量要远小于Ⅲ级岩体的数据量,若不处理这一问题而直接构建模型,则所得模型在小样本类别上的识别性能将会受到影响。因此,对训练集内样本量少的Ⅰ级岩体和Ⅳ级岩体数据,利用MDO过采样技术处理其训练集数据,使之与Ⅲ级岩体训练集样本数相等,即过采样后Ⅰ、Ⅳ级岩体的训练集数据量也为4524,以利于后续XGBoost算法构建岩体等级识别模型。

图10展示了Ⅰ级和Ⅲ级的训练集数据在过采样前后的数据特征,无论是均值还是标准差,过采样前后的特征差异均较小。这说明MDO算法在过采样过程中能够充分考虑到原始数据分布特征,新数据未破坏原始数据的分布结构,为后续模型训练提供了较好的数据基础。

#### 3.2 模型的超参数寻优与模型构建

在利用MDO算法过采样处理训练集中Ⅰ、Ⅳ级岩体的样本后,基于XGBoost算法拟合钻凿参数与岩体等级的关系。其中,训练模型时的输入特征为钻进速度、冲击压强、推进压强、旋转压强和旋转速度阀门开度等5个钻凿参数,模型的输出值为岩体等级(Ⅰ~Ⅳ级)。

XGBoost模型构建过程中具有较多的超参数,选取合理的超参数数值是构建高性能模型的基础。表3展示了XGBoost模型关键的超参数和取值范围,考虑到网格搜索算法的计算量极大,本文利用贝叶斯算法对各个超参数的取值进行寻优,设置贝叶

斯寻优的初始样本数量为 30,迭代次数为 500。在利用贝叶斯进行迭代寻优时,首先利用训练集中 80%的数据进行模型训练,将剩余 20%的数据作为

验证集,利用模型在验证集上的评估指标数值大小来评价每次迭代时超参数取值的优劣。

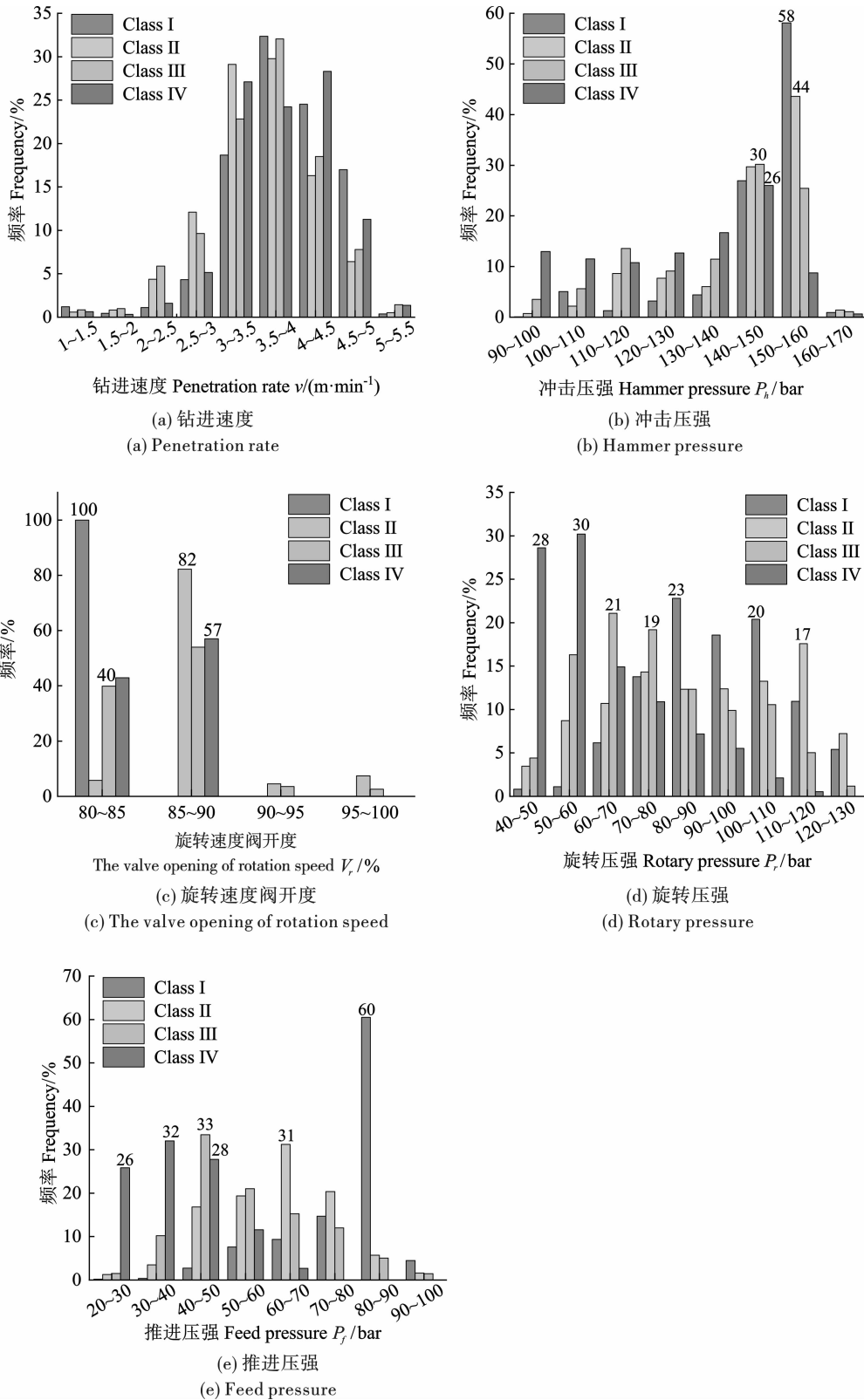


图 9 不同岩体条件下各钻凿参数数据分布

Fig. 9 Distribution of drilling parameters in different rock mass conditions

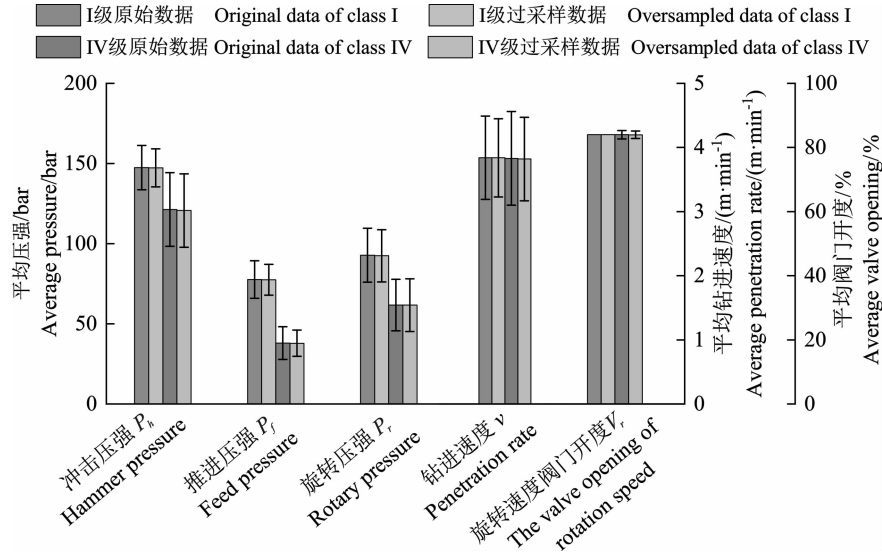


图 10 过采样前后钻凿数据的均值与标准差

Fig. 10 Mean and standard deviation of drilling data before and after oversampling

表 3 模型的超参数及其寻优范围

Table 3 Model hyperparameters and the search scope

超参数 Hyper-parameters	意义 Meaning	寻优范围 Optimization range
n_estimators	训练轮数, 控制决策树数量 The number of training rounds used to control the number of decision trees	0 ~ 300
learning_rate	学习率, 控制每棵树的贡献程度 The learning rate used to control the contribution level of each tree	0.01 ~ 0.5
min_child_weight	最小样本分割权重, 控制过拟合 Minimum sample segmentation weight used to control overfitting	1 ~ 10
max_depth	每棵树的深度, 控制树的复杂度 The maximum depth of each tree used to control complexity of tree	1 ~ 10
gamma	节点分裂所需的最小损失减少量 The minimum loss reduction required for node splitting	1 ~ 20
subsample	训练时的样本采样率, 用于避免过拟合 The sampling rate used to avoid overfitting during training process	0.1 ~ 1
colsample_bytree	训练决策树时的特征比例 The feature ratio during training	0.2 ~ 1
reg_alpha	$L_1$ 正则化项, 增加模型的稀疏性 $L_1$ regularization term used to increase the sparsity of model	0 ~ 20
reg_lambda	$L_2$ 正则化项, 防止过拟合 $L_2$ regularization term used to prevent overfitting	0 ~ 20

为防止训练次数过多造成过拟合, 在 XGBoost 模型中令参数 `early_stopping_rounds = 5` 设置早停机制。当指定的评估指标在连续 5 次的迭代中均没有改善时, 将提前停止训练。由于本文研究的问题属于为多分类任务, 因此将多类对数损失  $l$  作为评估指标, 其计算方法如式(1), 模型的识别准确率越高,  $l$  值越小。

$$l = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^K y_{ij} \log(p_{ij}) \quad (1)$$

式中:  $N$  为样本总数;  $K$  是类别总数;  $y_{ij}$  表示样本  $i$  对样本  $j$  的真实标签, 若样本  $i$  属于类别  $j$ , 则  $y_{ij}$  取值为 1, 否则取值为 0;  $p_{ij}$  是模型将样本  $i$  预测为类别  $j$  的概率。

图 11 为贝叶斯优选超参数时, 每次迭代所构建模型的  $l$  数值情况。由图可知, 贝叶斯算法在迭代初期训练出来的模型对数损失值较大, 且波动幅度明显; 当迭代次数超过 200 次后, 训练过程中偶尔会出现较大损失值, 但数值均未超过 0.8。

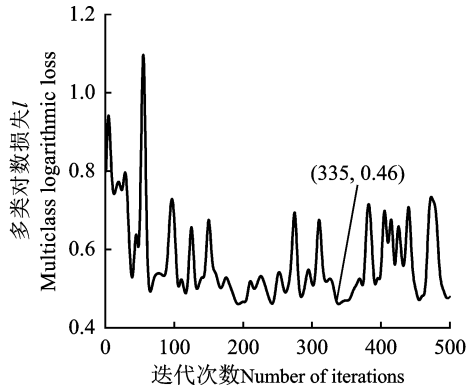


图 11 贝叶斯寻优过程中评估指标变化曲线  
Fig. 11 The change curve of evaluation index in the process of bayesian optimization

在贝叶斯算法进行第 335 次迭代时,所构建模型在验证集上的  $l$  最小,意味着模型在验证集上的识别性能最好,对应的超参数值如表 4 所示。将上述优选的超参数代入 XGBoost 算法中进行训练,构建出基于钻凿参数的岩体等级识别模型。

表 4 模型超参数的最优值

Table 4 The optimal value of the model hyperparameter

超参数 Hyper-parameters	最优值 Optimal value	超参数 Hyper-parameters	最优值 Optimal value
n_estimators	133	subsample	1
learning_rate	0.5	colsample_bytree	1
min_child_weight	1	reg_alpha	0
max_depth	10	reg_lambda	5.25
gamma	0.1		

### 3.3 模型评估

为评估所构建的 MDO-XGBoost 模型的识别性能,将之前所划分的测试集数据代入到模型中进行

测试评估,得到岩体等级识别模型在测试集上识别结果的混淆矩阵(如图 12)。图中对角线上数值为各类别预测正确的样本数量,模型整体识别率等于对角线数值和与矩阵所有数值之和之比。经过计算,模型在测试集上整体识别率达到了 80%,各类岩体的大部分样本均被准确识别。

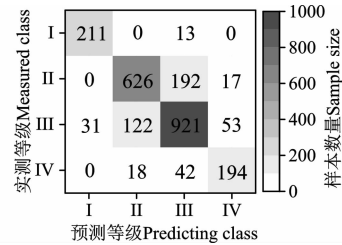


图 12 测试集预测结果的混淆矩阵

Fig. 12 Confusion matrix for test set prediction results

基于 SVM 算法、GBDT 算法和单纯的 XGBoost 算法对相同训练集进行模型构建,训练出 SVM 模型、GBDT 模型和未过采样的 XGBoost 模型,统计 SVM、GBDT、XGBoost 以及本文所提出的 MDO-XGBoost 四种模型在各等级岩体中的样本识别率(如表 5)。其中,SVM 模型的识别效果最差,其对 II 级岩体和 IV 级岩体的准确率小于 50%;GBDT 模型对 I ~ III 级岩体的识别率均在 70% 以上,对 III 级岩体的识别率还略高于 MDO-XGBoost 模型,但对于 IV 级岩体的识别率较低,说明其识别性能仍受样本不均衡的较大影响;MDO 和 XGBoost 的联合算法在 4 类岩体中的识别率均大于 70%,其中对 IV 级岩体的识别率达到了 76.38%,比 SVM、GBDT 分别提高了 34.65% 和 9.84%,比单纯的 XGBoost 算法提高了 5.74%。综上分析,本文模型能够在保持整体识别率 80% 的同时,一定程度上减小样本不均衡对小样本类识别效果的影响,改善模型的识别性能。

表 5 不同模型的识别率

Table 5 The recognition accuracy of different models

模型 Models	样本识别率 Sample identification rate/%			
	I 级 Class I	II 级 Class II	III 级 Class III	IV 级 Class IV
SVM	68.30	34.13	85.45	41.73
GBDT	93.30	70.89	85.27	66.54
XGBoost	93.90	73.85	80.20	70.64
MDO - XGBoost	94.20	74.97	81.72	76.38

特征重要度是评价各特征对所构建模型的贡献程度的指标,MDO-XGBoost 模型可根据每个特征在

所有树中对预测结果的平均增益来计算各特征的重要度,平均增益越大,意味着该特征在节点上的分裂

能力越强,对模型预测的贡献越大。图 13 展示了本模型各钻凿参数作为输入特征的重要程度,其中推进压强和旋转速度相较于其他参数更加重要,重要性评分分别为 0.32 和 0.30。

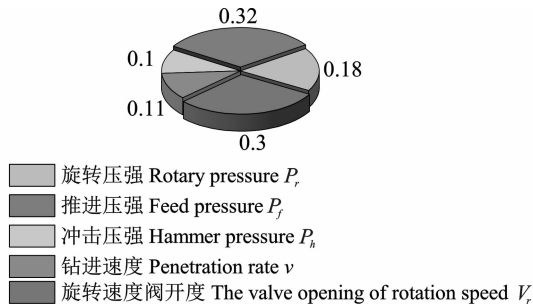


图 13 XGBoost 模型各钻凿参数的特征重要度

Fig. 13 Importance scoring of tunneling parameters in XGBoost model

由图 13 可知,钻进速度在现场数据中的重要性较低,结合 2.3 节的分析可知,钻进速度在现场钻凿作业的过程中,受到操作手人为控制因素的影响较大,因此,在基于现场实际钻凿参数的岩体等级识别

研究中,钻进速度的数值变化相对较小。

## 4 现场应用

依托于大巴山隧道工程项目,将所构建模型应用于现场钻爆施工。提取某次钻凿作业的钻凿数据作为输入,利用模型识别各钻孔处的岩体等级,然后利用 Kriging 插值法获取整个开挖断面的岩体等级分布。现场工作人员基于岩体等级识别结果,确定识别为 I 级和 II 级的岩体位置,在凿岩台车配套的炮孔设计软件 Bever Plan 中找到相应坐标,并在此位置适当加密炮孔,减小炮孔孔距,优化原始爆破方案。

图 14 为原始的 III 级岩体条件下炮孔设计,图 15 为利用 MDO-XGBoost 模型得到断面上岩体等级的识别结果。由图 15 可知,该断面岩体主要为 II 级岩体和 III 级岩体,与现场勘探情况吻合。由于原爆破方案是按照 III 级岩体进行设计的,为了兼顾不同等级岩体,在识别结果为 I 级和 II 级岩体的位置,适当加密炮孔,减小炮孔孔距,图 15 中白色炮孔即为在原方案基础上增添的炮孔。

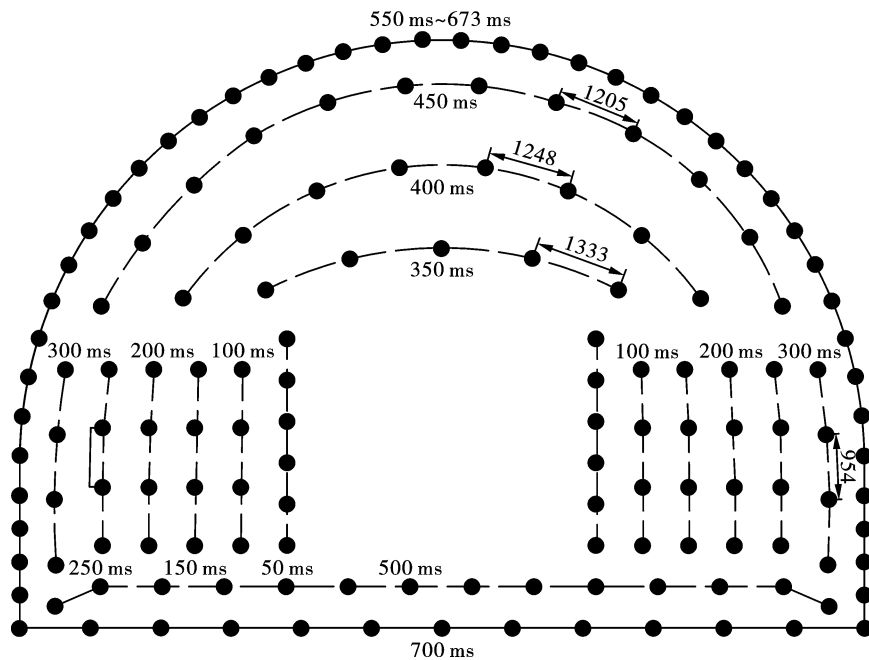


图 14 原始爆破方案(单位: mm)

Fig. 14 Original blasting plan(unit: mm)

将优化后的方案应用于现场,爆后断面情况如图 16 所示,整个断面较为平整,且可清晰看见半孔痕。图 17 展示了方案优化前后炮孔出渣所需时间与进尺深度(设计进尺 4~4.2 m),其中前四次爆破所用方案为原始方案,后四次爆破所用方案为改进后的方案。原始方案 4 个循环的平均出渣时间为

256 min、平均进尺深度为 3.325 m;与原始方案相比,改进爆破方案后的出渣时间整体上较少、平均时间为 226.5 min,进尺深度均大于 3.5 m,平均进尺达到 3.675 m。改进后的方案通过提高坚硬岩体位置的炸药单耗,降低了爆破的大块率,从而缩短了出渣时间,且提高了进尺深度,达到了更好的爆破效果。

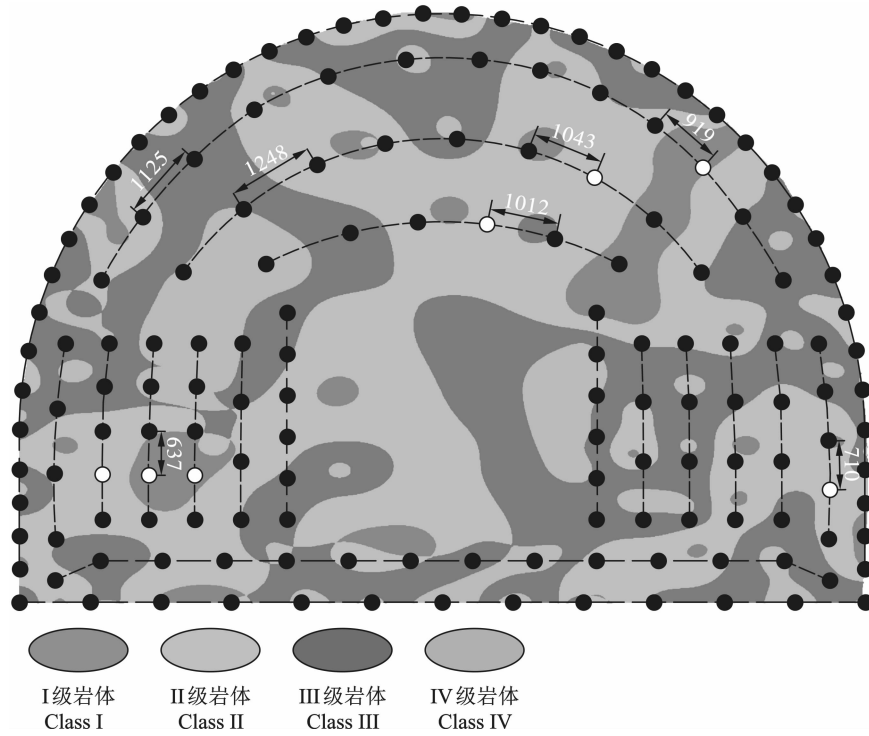


图 15 岩体等级识别情况及炮孔优化设计(单位:mm)

Fig. 15 Identification of rock mass classification and blast hole optimization(unit:mm)



图 16 现场爆破效果

Fig. 16 Field blasting effect

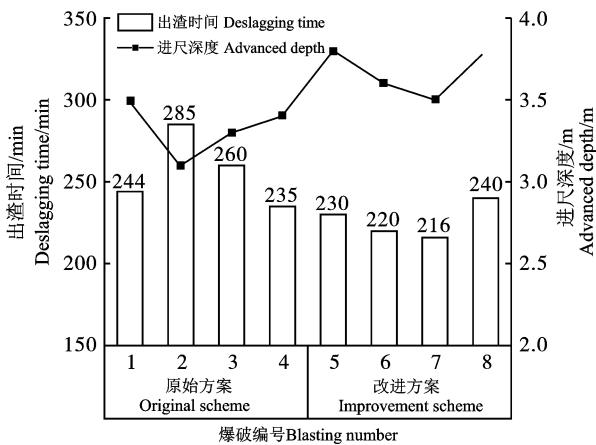


图 17 方案改进前后出渣时间与进尺深度

Fig. 17 Deslagging time and advanced depth before and after the scheme improvement

### 5 结论

研究依托重庆大巴山隧道工程项目,通过现场试验的方式采集到真实有效的钻凿参数和岩体特征,开展了样本不均衡情况下的钻凿参数与岩体等级关系研究,并以此优化了现场爆破方案,得到结论如下:

1)通过分析不同岩体条件下钻凿参数的数据分布特征,确定了与岩体等级相关的钻凿参数有推进压强、旋转压强、冲击压强、推进速度、旋转速度(阀门开度)。分析 XGBoost 岩体等级识别模型的特征重要度,进一步确定了推进压强和旋转速度对模型的重要度较高。

2)采用 MDO 算法处理现场所采集的样本比例不均衡的数据时,可以在保证数据均值和方差基本不变情况下,扩充了数据集,为后续岩体等级识别模型的构建提供良好数据基础。

3)提出了基于 MDO 过采样技术和 XGBoost 算法研究钻凿参数与岩体等级关系的方法,构建了整体识别率达到 80%的岩体等级识别模型,并提高了样本不均衡情况下对小样本类别的识别率,为隧道施工中优化爆破方案提供了有力支撑,优化后的爆破方案提高了进尺深度和出渣效率。基于钻凿参数的岩体等级识别模型有助于地质条件复杂的隧道施工中的智能化爆破建设。

## 参考文献 (References)

- [1] 岳中琦. 钻孔过程监测(DPM)对工程岩体质量评价方法的完善与提升[J]. 岩石力学与工程学报, 2014, 33(10):1977-1996.
- [1] YUE Zhong-qi. Drilling process monitoring for refining and upgrading rock mass quality classification methods [J]. Chinese Journal of Rock Mechanics and Engineering, 2014, 33(10):1977-1996. (in Chinese)
- [2] 李利平, 邹浩, 刘洪亮, 等. 钻爆法隧道智能建造研究现状与发展趋势[J]. 中国公路学报, 2024, 37(7):1-21.
- [2] LI Li-ping, ZOU Hao, LIU Hong-liang, et al. Research status and development trends in intelligent construction of drill-and-blast tunnels[J]. China Journal of Highway and Transport, 2024, 37(7):1-21. (in Chinese)
- [3] 王明年, 易文豪, 赵思光, 等. 基于钻进参数的隧道掌子面围岩单轴抗压强度和弹性模量解析[J]. 中国公路学报, 2024, 37(7):22-34.
- [3] WANG Ming-nian, YI Wen-hao, ZHAO Si-guang, et al. Analysis of uniaxial compressive strength and elasticity modulus of tunnel face surrounding rock based on drilling parameters[J]. China Journal of Highway and Transport, 2024, 37(7):22-34. (in Chinese)
- [4] 陈应显, 周萌. 基于智能岩性识别的炮孔装药量计算[J]. 爆破, 2022, 39(4):92-99.
- [4] CHEN Ying-xian, ZHOU Meng. Blasting charge calculation based on intelligent lithology identification[J]. Blasting, 2022, 39(4):92-99. (in Chinese)
- [5] 何川, 陈子全, 周子寒, 等. 基于机器学习的隧道支护体系智能化设计与评价方法[J]. 中国公路学报, 2023, 36(11):205-217.
- [5] HE Chuan, CHEN Zi-quan, ZHOU Zi-han, et al. Intelligent design and evaluation method of tunnel support system based on machine learning [J]. China Journal of Highway and Transport, 2023, 36(11):205-217. (in Chinese)
- [6] 李宁, 李骞, 宋玲. 基于回转切削的岩石力学参数获取新思路[J]. 岩石力学与工程学报, 2015, 34(2):323-329.
- [6] LI Ning, LI Qian, SONG Ling. Acquiring mechanical parameters of rock based on rotational cutting[J]. Chinese Journal of Rock Mechanics and Engineering, 2015, 34(2):323-329. (in Chinese)
- [7] 曹瑞琅, 王玉杰, 赵宇飞, 等. 基于钻进过程指数定量评价岩体完整性原位试验研究[J]. 岩土工程学报, 2021, 43(4):679-687.
- [7] CAO Rui-lang, WANG Yu-jie, ZHAO Fei-yu, et al. In-situ tests on quantitative evaluation of rock mass integrity based on drilling process index [J]. Chinese Journal of Geotechnical Engineering, 2021, 43(4):679-687. (in Chinese)
- [8] 易文豪, 王明年, 童建军, 等. 基于支持向量机的大断面岩质隧道掌子面围岩非均一性判别方法[J]. 中国铁道科学, 2021, 42(5):112-122.
- [8] YI Wen-hao, WANG Min-nian, TONG Jian-jun, et al. Inhomogeneity identification method for surrounding rock of large-section rock tunnel face based on support vector machine [J]. China Railway Science, 2021, 42(5):112-122. (in Chinese)
- [9] 张幼振, 张宁, 邵俊杰, 等. 基于钻进参数聚类的含煤地层岩性模糊识别[J]. 煤炭学报, 2019, 44(8):2328-2335.
- [9] ZHANG You-zhen, ZHANG Ning, SHAO Jun-jie, et al. Fuzzy identification of coal-bearing strata lithology based on drilling parameter clustering [J]. Journal of China Coal Society, 2019, 44(8):2328-2335. (in Chinese)
- [10] GHOSH R, GUSTAFSON A, SCHUNNESSON H. Development of a geological model for chargeability assessment of borehole using drill monitoring technique [J]. International Journal of Rock Mechanics and Mining Sciences, 2018, 109:9-18.
- [11] 王明年, 赵思光, 童建军, 等. 基于炮孔钻进参数的隧道掌子面围岩三维精细化分级方法[J]. 铁道学报, 2024, 46(10):163-173.
- [11] WANG Min-nian, ZHAO Si-guang, TONG Jian-jun, et al. Three-dimensional fine classification method for tunnel face surrounding rock based on blasting hole drilling parameters of drilling jumbo [J]. Journal of the China Railway Society, 2024, 46(10):163-173. (in Chinese)
- [12] 梁铭, 彭浩, 解威威, 等. 基于数字钻探与多尺度模型融合的隧道岩体完整性自动解译技术研究及应用[J]. 岩土工程学报, 2024, 46(2):396-405.
- [12] LIANG Ming, PENG Hao, XIE Wei-wei, et al. Research and application of automatic interpretation technology of tunnel rock mass integrity based on digital drilling and multi-scale model fusion [J]. Chinese Journal of Geotechnical Engineering, 2024, 46(2):396-405. (in Chinese)
- [13] 和郑翔, 卢才武, 居培, 等. 基于PCA-BP神经网络的随钻参数岩性智能感知方法研究[J]. 矿业研究与开发, 2022, 42(7):155-159.
- [13] HE Zheng-xiang, LU Cai-wu, JU Pei, et al. Research on lithology intelligent sensing method of drilling parameters based on PCA-BP neural network [J]. Mining Research and Development, 2022, 42(7):155-159. (in Chinese)
- [14] 姚宁平, 吴敏, 陈略峰, 等. 煤矿坑道钻进过程智能优化与控制技术[J]. 煤田地质与勘探, 2023, 51(9):

- 1-9.
- [14] YAO Ning-ping, WU Min, CHEN Lue-feng, et al. Intelligent optimization and control technology for drilling process of coal mine tunnels[J]. *Coal Geology and Exploration*, 2023, 51(9):1-9. (in Chinese)
- [15] ZHANG Qian-li, LIU Zhen-yu, TAN Jian-rong. Prediction of geological conditions for a tunnel boring machine using big operational data[J]. *Automation in Construction*, 2019, 100:73-83.
- [16] LIU Quan-sheng, WANG Xin-yu, HUANG Xing, et al. Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM drilling data[J]. *Tunnelling and Underground Space Technology*, 2020, 106:103595.
- [17] 王光宇, 宋建国, 徐飞, 等. 不平衡样本集随机森林岩性预测方法[J]. *石油地球物理勘探*, 2021, 56(4):679-687, 669.
- [17] WANG Guang-yu, SONG Jian-guo, XU Fei, et al. Random Forests lithology prediction method for imbalanced data sets [J]. *Oil Geophysical Prospecting*, 2021, 56(4):679-687, 669. (in Chinese)
- [18] BATISTA G E A P A, PRATI R C, MONARD M C. A study of the behavior of several methods for balancing machine learning training data[J]. *ACM SIGKDD Explorations Newsletter*, 2004, 6(1):20 - 29. DOI:10.1145/1007730.1007735
- [19] 朱梦琦, 朱合华, 王昕, 等. 基于集成 CART 算法的 TBM 掘进参数与围岩等级预测[J]. *岩石力学与工程学报*, 2020, 39(9):1860-1871.
- [19] ZHU Meng-qi, ZHU He-hua, WANG Xin, et al. Study on CART-based ensemble learning algorithms for predicting TBM tunneling parameters and classing surrounding rockmasses[J]. *Chinese Journal of Rock Mechanics and Engineering*, 2020, 39(9):1860-1871. (in Chinese)
- [20] YU Zhi-chao, WANG Zhi-zhang, ZENG Fan-cheng, et al. Volcanic lithology identification based on parameter-optimized GBDT algorithm: A case study in the Jilin Oilfield, Songliao Basin, NE China[J]. *Journal of Applied Geophysics*, 2021, 194:104443.
- [21] 张晨, 沈波, 刘桂卫. 浩吉铁路某长深隧道地应力三维反演技术研究[J]. *铁道工程学报*, 2022, 39(8):30-34.
- [21] ZHANG Chen, SHEN Bo, LIU Gui-wei. Research on the three-dimensional inversion technology of in-situ stress in a long and deep tunnel of Haolebaoji-Ji'an railway [J]. *Journal of Railway Engineering Society*, 2022, 39(8):30-34. (in Chinese)
- [22] 张凌凡, 陈忠辉, 周天白, 等. 基于梯度提升决策树的露天矿边坡多源信息融合与稳定性预测[J]. *煤炭学报*, 2020, 45(S1):173-180.
- [22] ZHANG Ling-fan, CHEN Zhong-hui, ZHOU Tian-bai, et al. Multi-source information fusion and stability prediction of slope based on gradient boosting decision tree [J]. *Journal of China Coal Society*, 2020, 45(S1):173-180. (in Chinese)
- [23] 韩启迪, 张小桐, 申维. 基于梯度提升决策树(GBDT)算法的岩性识别技术[J]. *矿物岩石地球化学通报*, 2018, 37(6):1173-1180.
- [23] HAN Qi-di, ZHANG Xiao-tong, SHEN Wei. Lithology identification technology based on gradient boosting decision tree(GBDT) algorithm[J]. *Bulletin of Mineralogy, Petrology and Geochemistry*, 2018, 37(6):1173-1180. (in Chinese)
- [24] ABDI L, HASHEMI S. To combat multi-class imbalanced problems by means of over-sampling techniques [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2016, 28(1):238-251.
- [25] CHEN Tian-qi, GUESTRIN C. XGBoost: A scalable tree boosting system [C] // *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2016:785-794.
- [26] 李山, 田仁飞, 刘涛. 应用 XGBoost 算法的随机缺失地震数据重建[J]. *石油地球物理勘探*, 2024, 59(5):965-975.
- [26] LI Shan, TIAN Ren-fei, LIU Tao. Reconstruction of randomly missing seismic data using XGBoost algorithm [J]. *Oil Geophysical Prospecting*, 2024, 59(5):965-975. (in Chinese)
- [27] 许江波, 侯鑫敏, 吴雄, 等. 基于 MIC-XGBoost-LSTM 模型的边坡位移预测研究[J]. *中国公路学报*, 2024, 37(10):38-48.
- [27] XU Jiang-bo, HOU Xin-min, WU Xiong, et al. Research on slope displacement prediction based on MIC-XGBoost-LSTM model [J]. *China Journal of Highway and Transport*, 2024, 37(10):38-48. (in Chinese)
- [28] 中华人民共和国住房和城乡建设部, 中华人民共和国国家质量监督检验检疫总局. GB/T 50218-2014 工程岩体分级标准:第4部分 岩体基本质量分级[S]. 北京:中国计划出版社, 2014:8-9.
- [28] Ministry of Housing and Urban-Rural Development of the People's Republic of China, General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China. GB/T 50218-2014 Standard for engineering classification of rock mass: Part 4 Classification of rock mass basic quality [S]. Beijing: China Planning Press, 2014:8-9. (in Chinese)