



文献 CSTR:
32001.14.11-6035.csd.2025.0100.zh



文献 DOI:
10.11922/11-6035.csd.2025.0100.zh
数据 DOI:
10.12262/lscd.Geography.2357

文献分类: 测绘科学技术

收稿日期: 2025-05-26
录用日期: 2025-10-24
发表日期: 2025-11-21

露天煤矿无人机多模态语义分割遥感数据集

崔朋志^{1,2}, 刘宝锋^{3*}, 王翀⁴, 李新鹏⁵, 张雯慧¹, 王迷军⁶,

刘吉祥¹, 李卓航¹, 李九龙¹, 丁颖³

1. 应急管理部信息研究院, 北京 100029
2. 辽宁工程技术大学, 测绘与地理科学学院, 辽宁阜新 123000
3. 咸阳职业技术学院, 陕西咸阳 712000
4. 华北科技学院, 计算机科学与工程学院, 河北廊坊 065201
5. 国能宝日希勒露天煤矿, 内蒙古呼伦贝尔 021000
6. 中安智讯(北京)信息科技有限公司, 北京 100029

摘要: 露天煤矿具有空间尺度大、地表形态复杂、动态要素多变等特征。传统人工巡检与单一监测手段存在效率低、盲区多、风险高等显著缺陷, 严重威胁煤矿生产安全。利用信息化技术手段, 及时、全面、精准地掌握矿区开采全过程对于安全生产具有重要意义。本研究利用无人机搭载激光雷达和五镜头倾斜相机进行数据采集, 构建了融合数字正射影像 (DOM)、数字表面模型 (DSM)、激光点云与三维倾斜模型的多模态遥感数据集, 并提取出 2992 张图像样本进行语义分割标注, 构建了首个露天煤矿多模态语义分割数据集, 数据量约 12.5 GB。通过 Unet、RandLA-Net 和 IKD-Net 模型分别从图像、点云及多模态融合角度进行训练与预测, 实验表明多模态融合策略显著提升分割精度: 图像模态 (U-Net) 准确率 92.24%, 点云模态 (RandLA-Net) 79.15%, 而多模态融合模型 (IKD-Net) 达到 94.9%, 据此建立了本数据集的性能评估基准。本数据集对推动露天矿无人化安全作业与可持续发展具有重要科研价值与工程应用前景。

关键词: 无人机; 露天煤矿; 遥感; 语义分割; 多模态

数据库(集)基本信息简介

数据库(集)名称	露天煤矿无人机多模态语义分割遥感数据集
数据通信作者	刘宝锋 (285463526@qq.com)
数据作者	崔朋志、刘宝锋、王翀、李新鹏、张雯慧、王迷军、刘吉祥、李卓航、李九龙、丁颖
数据时间范围	2022年6月
地理区域	内蒙古自治区呼伦贝尔市某大型露天矿
空间分辨率	0.51 m
数据量	12.5 GB
数据格式	*.tif, *.laz, *.osgb
数据服务系统网址	https://doi.org/10.12262/lscd.Geography.2357 https://scidb.cn/doi/10.12262/lscd.Geography.2357

*论文通信作者
刘宝锋: 285463526@qq.com

基金项目	国家自然科学基金（52274159）；应急管理部信息研究院基金（2023502）；咸阳市无人机测绘重点实验室和BIM数字建造重点实验室阶段性成果。
数据库（集）组成	本数据集包含无人机航测的成果数据，按照类型分为： （1）正射影像数据，以RGB三通道的GeoTif格式存储在DOM文件夹内； （2）数字表面模型数据，以单通道的GeoTif格式存储在DSM文件夹内； （3）倾斜摄影数据，以OSGB格式存储在OSGB文件夹内； （4）激光点云数据，以LAZ格式存储在LiDAR文件夹内，点的属性包含空间坐标、反射强度及RGB真彩色信息。 本数据集标注后建立的多模态语义分割数据集，按照模态存储在OPMMD下不同的文件夹内： （1）Image文件存放可见光图像，以RGB三通道的TIF格式存储； （2）Dsm文件夹存放高程数据，以单通道的TIF格式存储，其值为高程值； （3）LiDAR文件夹存放激光点云数据，以LAZ格式存储； （4）Mask文件夹存放标记后的数据，以单通道的TIF格式存储，其值为分类编号（0-5）。

引言

露天煤矿的智能化、信息化转型是实现国家能源安全与绿色矿山战略的关键。然而，其空间尺度大、地表复杂、动态多变的作业场景，导致传统人工巡检和单一监测手段存在效率低下、盲区多、风险高等问题。因此构建多维度矿山数据集是实现矿区信息化管控和安全高效开采的有效手段，更是响应国家“十四五”智慧矿山建设规划的重要实践。

近年来，面向矿区场景的专用数据集研发取得一定进展，露天煤矿分析从二维单模态向三维多源协同推进。其中，MABAC-FR-LG 模型通过无人机（UAV）航测构建数字高程模型（DEM），评估鄂尔多斯万利矿区沟谷侵蚀敏感性（GES），揭示坡度、高程及地裂缝距离等关键驱动因子^[1]；UAVI_Seg_Pytorch 数据集融合 RGB 与 DSM 高程数据，主要用大量样本解决地表覆盖分类问题^[2]。这些数据集为露天矿智能化监测与无人化作业奠定了初步数据基础，但仍普遍面临动态场景适应性弱、标注成本高等挑战，限制了其在复杂快速变化环境中的广泛应用。

本研究构建的露天煤矿遥感数据集，系统整合了 DOM、DSM、激光点云与三维倾斜模型^[3-7]等多模态时空数据，有效弥补了该领域高质量训练样本的不足。数据集通过创新性地融合空-地多源观测数据，突破单一数据模态在复杂场景解析中的局限性，为露天矿边坡^[5]稳定性分析、开采设备智能调度、自动驾驶路径规划等核心应用场景提供了关键数据支撑，同时对推动露天矿无人化安全作业与可持续发展具有重要科研价值与工程应用前景。

1 数据采集和处理方法

1.1 数据采集方法

数据采集于内蒙古自治区呼伦贝尔市某大型露天矿区，利用四川纵横无人机技术有限公司 CW-25E 垂直起降固定翼无人机飞行平台搭载激光雷达和倾斜相机于 2022 年 6 月 18 - 19 日完成数据采集，无人机以及载荷技术规格见表 1。

表 1 技术规格

Table 1 Technical specifications

无人机		载荷		
机身长度	2.18 m	激光雷达	最大测量距离	1845 m
翼展	4.35 m		精度	平面：0.01 m；垂直：0.02 m
最大起飞重量	31.58 kg		后处理 POS 系统姿态精度	航向：0.017°； 俯仰：≤0.005°
最大任务载荷	6.6 kg		定位系统	双频 GPS+双频 GLONASS+双频北斗
续航时间	150 min		存储、导出速度	100 Mbyte/s
巡航速度	70 km/h	倾斜相机	像素	2.1 亿
抗风能力	13.9-17.1 m/s		焦距	正射：40 mm；倾斜：56 mm
实用升限	6000 m		最小曝光间隔	0.6 s
最高起飞海拔	4000 m		供电方式	机载统一供电

首先在矿区四周边缘、采坑底部等部位布设像控点和标靶点，然后使用无人机五镜头倾斜相机对整个矿区进行全区域倾斜摄影，并利用无人机机载雷达采集全区域高精度点云，最后将倾斜影像点云与机载激光点云融合建模得到相应矿区数据集。

1.2 数据处理

外业获取的倾斜摄影测量成果，具体的处理步骤如下：首先基于外业采集的多视角影像、像控点及 POS 数据，通过空中三角测量解算影像外方位元素，具体流程包括：（1）利用 Context Capture 软件进行第一遍空三解算，基于 POS 数据提取连接点完成相对定向，建立影像间的空间关系。（2）导入地面控制点并通过人机交互刺点，结合同名连接点进行第二遍空三解算，执行光束法区域网平差优化，实现绝对定向，确保模型空间坐标精度；空三完成后需检查平差残差、控制点中误差等参数，验证成果可靠性。（3）通过密集匹配技术自动提取高精度三维点云，构建 TIN 三角网生成白模，并利用影像纹理映射技术赋予模型真实色彩，形成实景三维模型。（4）基于三维模型或点云生成 DOM、DSM 及 DEM，其中 DEM 需通过滤波算法滤除非地面点并经插值优化。

机载激光雷达（LiDAR）数据，处理流程侧重于点云分类与地形建模：（1）预处理阶段，通过 TerraSolid 等软件滤除噪声点，基于激光点云的空间分布、回波特性等参数进行自动分类，分离地面点、植被点及建筑物点。（2）人工干预阶段，结合正射影像对自动分类结果进行精细化调整，修正高程突变区域的误分类点。（3）基于分类后的地面点云插值生成高精度 DEM，同时利用原始点云生成包含地物顶部的 DSM，结合倾斜影像补充地物纹理细节，通过独立建模与融合建模试验提升数据成果完整性。

2 数据样本描述

2.1 多模态数据

本数据集包含数字正射影像（DOM）、数字表面模型（DMS）、激光点云和三维倾斜模型数据。

DOM 以 RGB 三通道的 TIFF 格式保存，分辨率达亚米级，可视化如图 1(a)；DSM 采用 16 位深度单通道 TIFF 格式，精确记录地表高程信息，可视化如图 1(b)；激光点云数据以 LAZ 格式封装，完整保留三维空间坐标、反射强度及 RGB 真彩色信息，可视化如图 1(c)；倾斜摄影模型则选用 OSGB 开放格式，支持多细节层次（LOD）的三维可视化，可视化如图 1(d)。

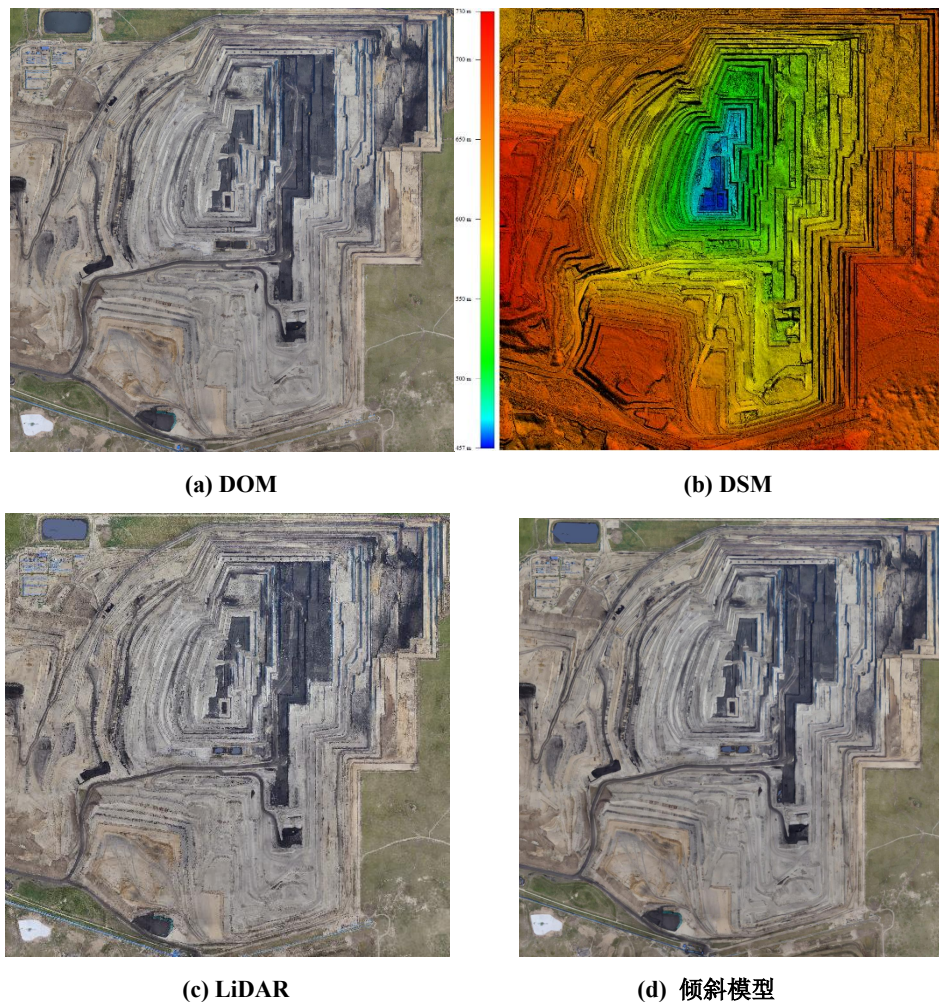


图 1 多模态数据

Figure 1 Images of multimodal data

基于上述数据，我们构建了首个露天矿多模态数据集（Open-Pit Mine Multimodal Dataset, OPMMD），有机融合了可见光影像（RGB）、数字地表模型（DSM）与激光雷达点云（LiDAR Point Cloud）三种模态。如表 2 所示，OPMMD 在关键指标上均处于领先地位：其覆盖空间范围达 24 km²，且模态完整性最高（同时包含 RGB、DSM 和 LiDAR）。此外，OPMMD 是国内外首个专门适用于露天矿场景的多模态语义分割数据集。

表 2 与语义分割代表性数据集的比较

Table 2 Comparison with the representative datasets for semantic segmentation

类别	名称	时间	空间范围	分辨率	模态	获取方式
室内场 景	S3DIS ^[8]	2017	6×10 ³ m ²	-	RGB-D	Matterport 相机
	ScanNet ^[9]	2017	1.1×10 ⁵ m ²	-	RGB-D	RGB-D 相机

类别	名称	时间	空间范围	分辨率	模态	获取方式
室外道路	Semantic3D ^[10]	2017	-	-	LiDAR	移动式激光扫描系统
	Toronto-3D ^[11]	2020	1×10 ³ m	-	LiDAR	移动式激光扫描系统
城市	Campus3D ^[12]	2020	1.6×10 ⁶ m ²	-	LiDAR	无人机摄影测量
	SensatUrban ^[13]	2020	7.6×10 ⁶ m ²	0.1m	LiDAR	无人机摄影测量
	Potsdam ^[14]	2014	3×10 ⁶ m ²	0.05m	RGB、DSM、LiDAR	无人机摄影测量
	Vaihingen ^[14]	2014	1.3×10 ⁶ m ²	0.1m	RGB、DSM、LiDAR	无人机摄影测量
	N3C-California ^[15]	2024	7.2×10 ⁸ m ²	0.6m	RGB、DSM、LiDAR	航空摄影测量
	GRSS ^[16]	2018	5×10 ⁶ m ²	0.5m	RGB、DSM、LiDAR	航空摄影测量
矿区	OPMMD	2025	2.4×10 ⁷ m ²	0.51m	RGB、DSM、LiDAR	无人机摄影测量

注：“-”表示该参数在相应的参考文件中未说明。

2.2 语义类别标签

本数据集分别采用 LabelMe 和 CloudCompare 软件对可见光影像和激光点云数据进行标注，语义类别划分基于以下 3 个原则：1) 每个类别应具有行业或商业价值，例如道路识别、地图构建等。2) 每个类别应具有明确且无歧义的语义含义。3) 不同类别在几何结构或外观上应有显著差异。为此，确定了 6 种语义类别，具体定义如表 3。

表 3 数据集分类定义

Table 3 Class definitions of the dataset

序号	类别	定义
1	地表	裸地、绿地等
2	道路	运输主干路、联络道路、工作面道路、排土场道路、铁路等
3	边坡	采矿场边坡、排土场边坡
4	车辆	电铲、矿用卡车、装载机、洒水车、推土车、加油车、指挥车等
5	建筑物	办公楼、简易房、洗选煤车间、煤仓、车库、水泵房等
6	水系	地表径流、积水坑等

每个像素和每个点都被标注为 6 种语义类别之一，所有标注都经过人工交叉检查，以确保一致性和高质量。图 2 展示了标注后的部分数据。

2.3 数据统计

与大多数语义分割数据集类似，将影像、点云、高程等数据划分为一系列大小相等的瓦片，以保证本数据能够输入 GPU 中进行训练和测试。具体而言，按照 512×512 像素对数据进行瓦片划分，从而得到尺寸为 512×512 分辨率的三通道图像瓦片和单通道的高程瓦片以及相应的点云瓦片，并实现坐标对齐，共得到瓦片数据 2992 张，并按照 8:1:1 的比例得到训练、测试和验证数据集的瓦片数量分别为 2392、300、300。数据的具体类别如图 3。

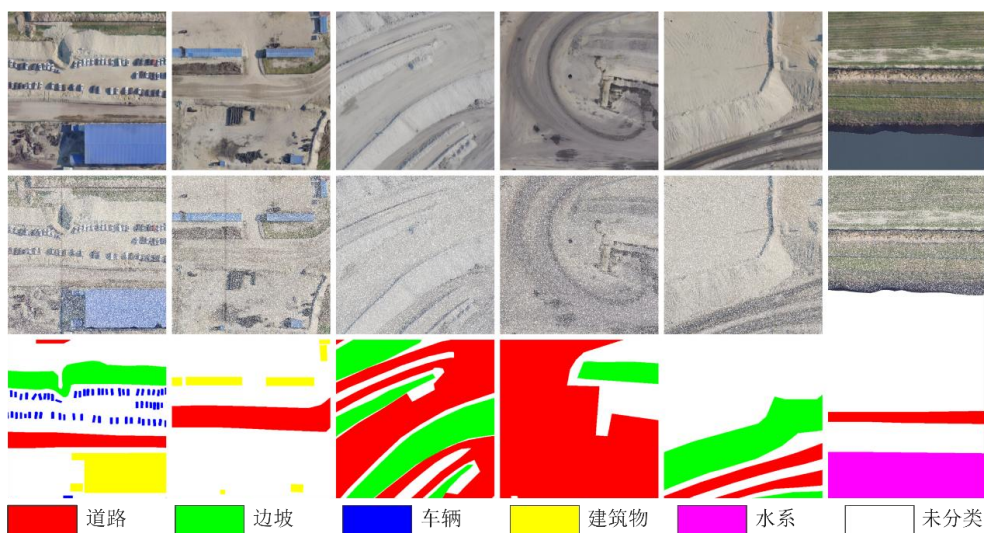


图2 数据集中影像、点云和语义标签瓦片可视化

Figure 2 Visualization of images, point clouds and semantic label files in the dataset

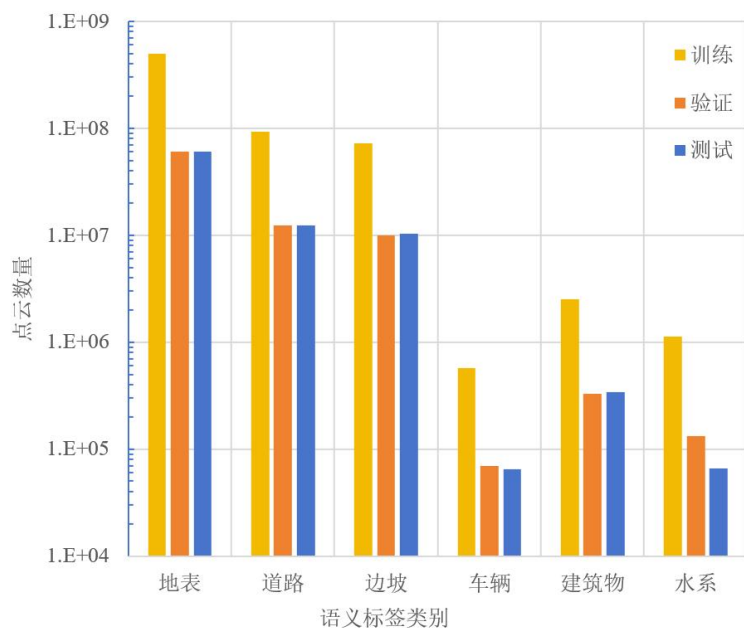


图3 数据集中各类别像素统计

Figure 3 Pixel statistics of each class in the dataset

在语义分割任务中，地表一般是作为背景不参与识别。从统计图中可以看出，道路、边坡是主要的语义类别，占据了总数的 95%以上，车辆和建筑物的合计占比不足 1%，这说明在真实场景中各类别分布极度不平衡，这为基于本数据集的语义分割实验带来了严峻的挑战。

3 数据质量控制和评估

3.1 数据质量控制

为保证数字正射影像（DOM）、数字表面模型（DSM）、激光点云、三维倾斜模型数据以及构

建的多模态遥感数据集的数据质量，本数据集从三个方面进行控制：（1）影像数据：通过实时检查影像清晰度、验证航向/旁向重叠率及激光点云密度，确保原始数据完整性。（2）航测数据处理：利用空三解算、点云滤波及坐标系对齐消除噪声与偏差。（3）为了确保标注质量，所有像素和点在第一轮由两位专业人员独立标注，第二轮则进行交叉验证，数据标注后进行了人工检查。

3.2 数据评估

为验证本样本集的有效性，在统一的硬件平台下，采用相同超参数（初始学习率、批次大小、优化器等），分别采用 UNet^[17]、RandLA-Net^[18]、IKD-Net^[15]等深度学习算法进行多分类语义识别模型训练和测试。Unet 是基于对称编解码结构的经典图像分割网络，通过跳跃连接融合浅层细节与深层语义信息。RandLA-Net 是针对大规模点云的高效语义分割网络，采用随机采样和局部特征聚合降低计算开销。IKD-Net 网络通过融合航空图像和 LiDAR 点云数据，旨在解决多模态数据融合中的不平衡问题，从而提高了地表覆盖分类的准确性。为了方便对该多模态数据集进行评估，本文采用语义分割领域平均准确率（mAcc）、总体准确率（OA）、Kappa 系数、交并比（IoU）等通用的评价指标进行评估。各模型在该数据上的表现如表 4 和图 4 所示。

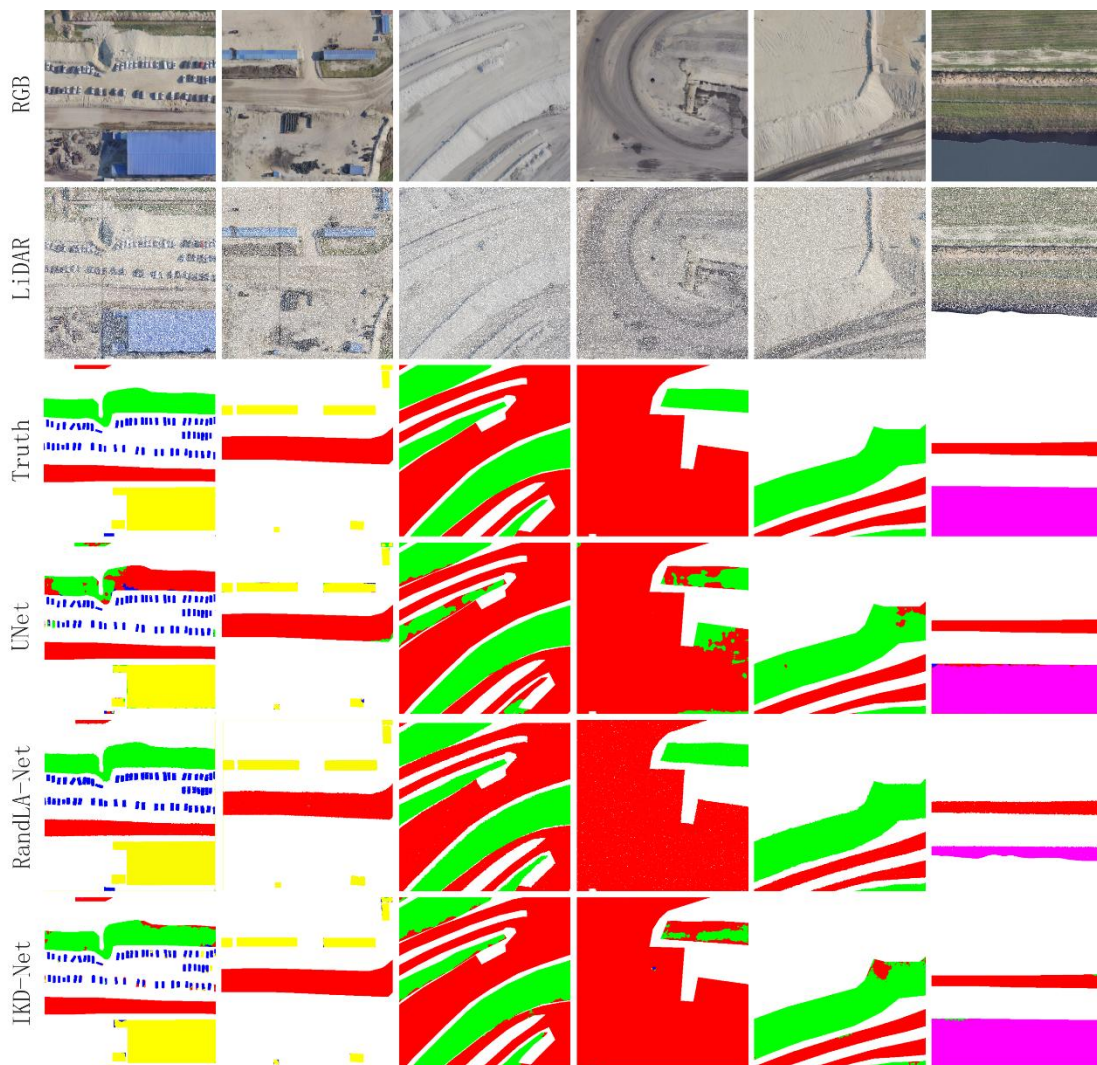


图 4 语义分割结果可视化

Figure 4 Visualization of semantic segmentation results

表 4 数据集分类定义

Table 4 Class definitions of the dataset

模型	输入	mAcc	OA	Kappa	IoU					
					道路	边坡	车辆	建筑	水系	平均
UNet	RGB	92.24	96.27	92.50	93.65	90.38	76.22	91.80	87.39	87.89
RandLA-Net	LiDAR	79.15	96.76	93.79	95.05	94.40	31.64	69.96	57.42	69.69
IKD-Net	RGB+LiDAR	94.90	97.52	95.00	96.11	94.05	62.98	94.36	77.86	85.07

表 4 中，平均准确率（mAcc）、总体准确率（OA）、Kappa 系数、交并比（IoU）和平均交并比（MioU）的定义和计算公式如下式（1）—（7）所示，式中 TP (True Positive)表示被分类为正样本，实际也是正样本，也就是正确预测的像素数；FP (False Positive) 表示被分类为正样本，但实际是负样本，也就是错误预测为正样本的像素数；FN (False Negative)表示被分类为负样本，但实际上是正样本，也就是错误预测为负样本的像素数；TN (True Negative)表示被分类为负样本，实际也是负样本，也就是正确预测负样本的像素数。

准确率表示为所有预测正确的结果占总预测值的比重，该指标用来评估全局准确率，当数据样本不平衡时，该指标不能客观评估模型，其表达式如式(1)：

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

总体准确率（Overall Accuracy, OA）表示所有预测正确的结果占总预测值的比重，侧重全局的整体正确率，其表达式如式(2)：

$$OA = \frac{\sum_{i=1}^n (TP_i + TN_i)}{\sum_{i=1}^n (TP_i + TN_i + FP_i + FN_i)} \quad (2)$$

平均准确率（Mean Accuracy, mAcc）表示各类别的准确率的算术平均值，避免了大类别主导的问题，适用于类别分类不均衡的任务，其表达式如式(3)：

$$mAcc = \frac{1}{n} \sum_{i=1}^n \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \quad (3)$$

Kappa 系数是一个用于一致性检验的指标，kappa 计算结果为[-1,1]，但通常 kappa 是落在[0,1]间。

$$Kappa = \frac{OA - p_e}{1 - p_e} \quad (4)$$

其中，

$$p_e = \frac{\sum_{i=1}^n (TP_i + FP_i)(TP_i + FN_i)}{\left(\sum_{i=1}^n (TP_i + TN_i + FP_i + FN_i) \right)^2} \quad (5)$$

交并比表示单个类别的预测区域与真实区域的重叠比例，其表达式如式(6)：

$$IoU = \frac{X \cap Y}{X \cup Y} = \frac{TP}{TP + FN + FP} \quad (6)$$

平均交并比表示所有类别的 IoU 的平均值，其表达式如式(7)：

$$MIoU = \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{TP_i + FN_i + FP_i} \quad (7)$$

4 数据价值

本数据集整合了数字正射影像（DOM）、数字表面模型（DSM）、激光点云和三维倾斜模型数据，旨在有效弥补当前露天煤矿遥感 CV 领域在多样化训练样本方面的不足。本数据集的应用价值：

（1）为复杂场景下的边坡、车辆等目标识别、高精度语义分割任务提供多模态数据支撑。（2）通过多源时空数据融合，支撑矿区数字孪生建模。（3）本数据可深度赋能露天矿智能化建设，特别是在无人采矿路径规划、自动驾驶系统等关键技术环节，为算法模型训练与验证提供真实场景数据基础。本数据资源的发布，对于推动智慧矿山理论方法创新与工程实践具有重要的科研价值与工程实践意义。

5 数据使用方法和建议

对于原始数据中的影像和倾斜模型，可使用 Envi、GlobalMapper、ArcGIS Pro 等软件进行查看。对于已标注数据，可直接通过 Python 中的 GDAL、laspy、cv2、PIL 等模块进行读取，再输出到深度学习模型中进行训练。

作者分工职责

崔朋志（1983—），男，河北正定人，硕士研究生，高级工程师，研究方向为煤矿三维可视化。主要承担工作：数据处理、算法编写、论文撰写。

刘宝锋（1979—），男，陕西咸阳人，硕士研究生，教授，研究方向为测绘地理信息。主要承担工作：数据处理、数据质量控制、论文撰写。

王翀（1974—），男，辽宁昌图人，博士研究生，高级工程师，研究方向为安全生产信息化技术。主要承担工作：算法分析、数据挖掘。

李新鹏（1993—），男，河南洛阳人，硕士研究生，工程师，研究方向为煤矿采矿。主要承担工作：数据采集、清洗和标注。

张雯慧（1996—），女，江西赣州人，硕士研究生，工程师，研究方向为煤矿三维可视化。主要承担工作：数据标注。

王迷军（1972—），男，内蒙古乌兰察布人，博士研究生，高级工程师，研究方向为煤矿信息化。主要承担工作：数据标注。

刘吉祥（1994—），女，山东日照人，硕士研究生，助理工程师，研究方向为煤矿三维地质建模。主要承担工作：数据标注。

李卓航（1997—），女，北京房山人，硕士研究生，助理工程师，研究方向为煤矿三维可视化。主要承担工作：数据标注。

李九龙（1991—），男，山东菏泽人，硕士研究生，中级工程师，研究方向为煤矿三维可视化。主要承担工作：数据标注。

丁颖（1979—），女，陕西铜川人，硕士研究生，讲师，研究方向为中英文翻译。主要承担工作：资料收集、论文中英文翻译。

参考文献

- [1] 万利矿区沟谷侵蚀敏感性评估与沟头前进速率研究：采矿扰动下的生态风险量化[EB/OL]. (2025-06-28) [2025-11-05]. <https://www.ebiotrade.com/newsf/2025-6/20250628073043627.htm>.
- [2] UAVI_Seg_Pytorch—无人机视角下的地面物体精细化分割库[EB/OL]. (2024-06-22) [2025-11-05]. https://blog.csdn.net/gitblog_00054/article/details/139876646.
- [3] 刘宝锋, 颀昕宇. 基于无人机倾斜摄影测量技术空地融合建模方法研究[J]. 微型电脑应用, 2024, 40(1): 59–61, 68. DOI:10.3969/j.issn.1007-757X.2024.01.016. [LIU B F, JIE X Y. Research on modeling method of air ground fusion based on UAV oblique photogrammetry technology[J]. Microcomputer Applications, 2024, 40(1): 59–61, 68. DOI:10.3969/j.issn.1007-757X.2024.01.016.]
- [4] 张书航, 李旺民, 庞尹宁, 等. 基于无人机倾斜摄影和激光扫描的离岸海岛高精度实景三维重建[J]. 测绘通报, 2024(3): 25–30. DOI: 10.13474/j.cnki.11-2246.2024.0305. [ZHANG S H, LI W M, PANG Y N, et al. Reconstruction of high-precision real 3D data of offshore islands based on drone tilt photography and laser scanning[J]. Bulletin of Surveying and Mapping, 2024(3): 25–30. DOI: 10.13474/j.cnki.11-2246.2024.0305.]
- [5] 李虹江, 于昕左, 马佳, 等. 多源遥感数据融合的高陡边坡危岩体信息提取[J/OL]. 地质科技通报, 1-16 [2025-05-19]. <https://doi.org/10.19509/j.cnki.dzkq.tb20230695>. [LI H J, Yu X Z, Ma J, et al. Extraction of Dangerous Rock Mass Information from High and Steep Slope by Multi source Remote Sensing Data Fusion [J/OL]. Geological Science and Technology Bulletin, 1-16 [205-05-19] <https://doi.org/10.19509/j.cnki.dzkq.tb20230695>.]
- [6] 胡晓静, 刘钰, 陈延博, 等. 融合多源数据的城市三维建模方法探索[J]. 测绘通报, 2023(11): 128–131. DOI: 10.13474/j.cnki.11-2246.2023.0340. [HU X J, LIU Y, CHEN Y B, et al. Research on 3D modeling method of city with multi-source data fusion[J]. Bulletin of Surveying and Mapping, 2023(11): 128–131. DOI: 10.13474/j.cnki.11-2246.2023.0340.]
- [7] 刘宝锋. 基于无人机倾斜摄影测量的三维可视化建模研究[J]. 自动化与仪器仪表, 2023(5): 304–309. DOI: 10.14016/j.cnki.1001-9227.2023.05.304. [LIU B F. Research on 3D visual Modeling for tilt photogrammetry of UAV[J]. Automation & Instrumentation, 2023(5): 304–309. DOI: 10.14016/j.cnki.1001-9227.2023.05.304.]
- [8] ARMENI I, SAX S, ZAMIR A R, et al. Joint 2D-3D-Semantic Data for Indoor Scene Understanding [J]. 2017. DOI:10.48550/arXiv.1702.01105.
- [9] DAI A, CHANG A X, SAVVA M, et al. ScanNet: richly-annotated 3D reconstructions of indoor scenes[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI,

- USA. IEEE, 2017: 2432–2443. DOI: 10.1109/CVPR.2017.261.
- [10] HACKEL T, SAVINOV N, LADICKY L, et al. semantic3d.net: a new large-scale point cloud classification benchmark[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2017, IV-1/W1: 91–98. DOI: 10.5194/isprs-annals-iv-1-w1-91-2017.
- [11] TAN W K, QIN N N, MA L F, et al. Toronto-3D: a large-scale mobile LiDAR dataset for semantic segmentation of urban roadways[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle, WA, USA. IEEE, 2020: 797–806. DOI: 10.1109/CVPRW50498.2020.00109.
- [12] LI X K, LI C S, TONG Z K, et al. Campus3D: a photogrammetry point cloud benchmark for hierarchical understanding of outdoor scene[EB/OL]. 2020: 2008.04968. <https://arxiv.org/abs/2008.04968v1>.
- [13] HU Q Y, YANG B, KHALID S, et al. SensatUrban: learning semantics from urban-scale photogrammetric point clouds[J]. International Journal of Computer Vision, 2022, 130(2): 316–343. DOI: 10.1007/s11263-021-01554-9.
- [14] ROTTENSTEINER F, SOHN G, GERKE M, et al. ISPRS semantic labeling contest [J]. ISPRS: Leopoldshöhe, Germany, 2014, 1(4): 4. DOI:10.13140/2.1.3570.944.
- [15] WANG Y M, WAN Y, ZHANG Y J, et al. Imbalance knowledge-driven multi-modal network for land-cover semantic segmentation using aerial images and LiDAR point clouds[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2023, 202: 385–404. DOI: 10.1016/j.isprsjprs.2023.06.014.
- [16] XU Y H, DU B, ZHANG L P, et al. Advanced multi-sensor optical remote sensing for urban land use and land cover classification: outcome of the 2018 IEEE GRSS data fusion contest[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(6): 1709–1724. DOI:10.1109/JSTARS.2019.2911113.
- [17] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Cham: Springer International Publishing, 2015: 234–241. DOI: 10.1007/978-3-319-24574-4_28.
- [18] HU Q Y, YANG B, XIE L H, et al. RandLA-net: efficient semantic segmentation of large-scale point clouds[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA. IEEE, 2020: 11105–11114. DOI: 10.1109/cvpr42600.2020.01112.

论文引用格式

崔朋志, 刘宝锋, 王翀, 等. 露天煤矿无人机多模态语义分割遥感数据集[J/OL]. 中国科学数据, 2025, 10(4). (2025-11-21). DOI:10.11922/11-6035.csd.2025.0100.zh.

数据引用格式

崔朋志, 刘宝锋, 王翀, 等. 露天煤矿无人机多模态语义分割遥感数据集[DS/OL]. 中国科学院地球环境研究所黄土科学数据中心, 2025. (2025-10-29). DOI:10.12262/lscd.Geography.2357.

A dataset of multi-modal UAV remote sensing data for semantic segmentation of open-pit coal mines

CUI Pengzhi^{1,2}, LIU Baofeng^{3*}, WANG Chong⁴, LI Xinpeng⁵, ZHANG Wenhui¹,
WANG Mijun⁶, LIU Jixiang¹, LI Zhuohang¹, LI JiuLong¹, DING Ying³

1. Information Research Institute, Ministry of Emergency Management, Beijing 100029, P.R. China

2. College of Mapping and Geographical Science, Liaoning Technical University, Fuxin 123000, P.R. China

3. Xianyang Polytechnic Institute, Xianyang 712000, P.R. China

4. College of Computer, North China Institute of Science and Technology, Langfang 065201, P.R. China

5. National Energy Group Baorixile Open-Pit Coal Mine, Hulunbuir 021000, P.R. China

6. China Safety Intelligence Information Technology Co., Ltd, Beijing 100029, P.R. China

*Email: 285463526@qq.com

Abstract: Open-pit coal mines are characterized by their large spatial scale, complex surface morphology, and highly variable dynamic elements. Traditional manual inspection and single-source monitoring methods suffer from low efficiency, multiple blind spots, and high operational risks, posing serious threats to production safety. Leveraging advanced information technologies to obtain timely, comprehensive, and accurate insights into the entire mining process is of great significance for safety operations. This study utilized unmanned aerial vehicles equipped with a LiDAR system and a five-lens oblique cameras for data acquisition, resulting in a multi-modal remote sensing dataset that integrates digital orthophoto maps (DOMs), digital surface models (DSMs), LiDAR point clouds, and 3D oblique models. 2,992 image samples were extracted and annotated for semantic segmentation. Based on this, the first multi-modal semantic segmentation dataset for open-pit coal mines was constructed, with a data volume of approximately 12.5 GB. Through training and prediction using the U-Net, RandLA-Net, and IKD-Net models—representing the image, point cloud, and multimodal fusion modalities, respectively—experiments demonstrate that the multimodal fusion strategy significantly improves segmentation accuracy: 92.24% for the image modality (U-Net), 79.15% for the point cloud modality (RandLA-Net), and 94.9% for the multimodal fusion model (IKD-Net). This dataset provides a valuable benchmark for open-pit mining research and holds significant scientific and engineering potential for advancing unmanned, safe, and sustainable operations in open-pit mines.

Keywords: UAV; open-pit coal mine; remote sensing; semantic segmentation; multi-modal

Dataset Profile

Title	A dataset of multi-modal UAV remote sensing data for semantic segmentation of open-pit coal mines
Data corresponding author	LIU Baofeng (285463526@qq.com)
Data authors	CUI Pengzhi, LIU Baofeng, WANG Chong, LI Xinpeng, ZHANG Wenhui, WANG Mijun, LIU Jixiang, LI Zhuohang, LI JiuLong, DING Ying

Time range	June, 2022
Geographical scope	A certain large-scale open-pit mine in Hulunbuir City, Inner Mongolia Autonomous Region, China
Spatial resolution	0.51 m
Data volume	12.5 GB
Data format	*.tif, *.laz, *.osgb
Data service system	https://doi.org/10.12262/lscd.Geography.2357 https://scidb.cn/doi/10.12262/lscd.Geography.2357
Source(s) of funding	National Natural Science Foundation of China (52274159); Institute-level Scientific Research Project of the Information Research Institute, Ministry of Emergency Management (2023502); Phased Achievements of Xianyang UAV Surveying and Mapping Key Laboratory.
Dataset composition	This dataset contains the results of UAV aerial survey, categorized by data type as follows: (1) Digital Orthophoto Data (DOM): Stored in RGB three-channel GeoTIFF format within the DOM folder; (2) Digital Surface Model (DSM): Stored in single-channel GeoTIFF format within the DSM folder; (3) Oblique Photography Data: Stored in OSGB format within the OSGB folder; (4) LiDAR Point Cloud Data: Stored in LAZ format within the LiDAR folder, and the attributes of points include spatial coordinates, reflection intensity and RGB true color information; The annotated multi-modal semantic segmentation dataset derived from this dataset is organized by modality within subfolders under OPMMD: (1) The Image folder contains visible-light imagery stored in RGB three-channel TIFF format; (2) The Dsm folder holds elevation data stored as single-channel TIFF files, where pixel values represent elevation; (3) The LiDAR folder includes LiDAR point cloud data stored in LAZ format; (4) The Mask folder contains annotated ground-truth stored as single-channel TIFF files, where pixel values indicate classification labels ranging from 0 to 5.