



文献 CSTR:

32001.14.11-6035.csd.2025.0028.zh



文献 DOI:

10.11922/11-6035.csd.2025.0028.zh

数据 DOI:

10.57760/sciencedb.21999

文献分类: 信息科学与系统科学

收稿日期: 2025-03-13

录用日期: 2025-05-07

发表日期: 2025-12-29

* 论文通信作者

路洋: luyang20000520@163.com

矿井遮挡人员跟踪数据集

高丽铄¹, 路洋^{2*}, 董立红², 叶鸥², 韩冷雨², 张立智², 郝晨阳²

1. 西安外事学院, 西安 710077

2. 西安科技大学, 西安 710054

摘要: 矿井作业环境复杂特殊, 给人员目标的检测与跟踪带来诸多挑战。作业人员通常穿着统一工服, 外观特征高度相似, 加之环境光照不足、图像模糊、背景灰暗单一, 显著降低了视觉识别的准确性。在矿井典型场景中, 人员密集、遮挡频发, 进一步增加了跟踪算法的难度。在深度学习模型训练中, 数据集的质量往往直接决定模型的性能。然而, 目前国内外尚缺乏公开、标准化的矿井人员图像数据集, 现有研究多依赖小规模、非规范的私有数据, 难以支撑算法的横向对比与推广应用。本研究通过使用煤矿用本安型执法记录仪对矿井环境中的遮挡人员现场进行拍摄, 整理构建了具有统一标准、大规模、高质量的矿井遮挡人员目标跟踪数据集。本数据集总共包含 14740 张图像, 总数据量为 38.1 GB, 囊括了三种矿井典型人员出现遮挡的场景(下井口出入口闸机(exit_gate)、井底硐室(well_chamber)和井下工作面(downhole_face))。本数据集的建立可有效缓解矿井遮挡人员目标跟踪数据资源不足的问题, 具有较高应用价值, 为相关研究提供了必要的数据支持。

关键词: 矿井; 目标跟踪; 遮挡人员; 深度学习; 数据集

数据库(集)基本信息简介

数据库(集)名称	矿井遮挡人员跟踪数据集
数据通信作者	路洋(luyang20000520@163.com)
数据作者	高丽铄、路洋、董立红、叶鸥、韩冷雨、张立智、郝晨阳
数据时间范围	2022–2024年
地理区域	陕西彬州
数据量	38.1 GB
数据格式	*.png, *.txt
数据服务系统网址	https://doi.org/10.57760/sciencedb.21999
基金项目	国家自然科学基金(62303375)
数据库(集)组成	本数据集由28个数据段组成, 每个数据段包含3部分: (1) det.txt文件中保存由目标检测算法生成的检测结果, 字段信息包含对应数据段中的图像帧编号、运动轨迹ID号、目标位置坐标、目标框长宽、范围标志值; (2) gt.txt文件中包含每一帧图像的目标跟踪标注数据, 标注信息包含对应数据段中每间隔33.3 ms所处时刻的图像帧编号、运动轨迹ID号、目标位置坐标、目标框长宽、目标遮挡情况、类别ID、可见度比例; (3) img1文件夹存放的是视频序列中的每一帧图像数据。

引言

矿井的作业环境独特，目标检测与跟踪面临一系列特有的挑战。首先，矿井作业人员的着装通常为统一的工作服，且大部分外观特征相似，这使得基于视觉特征的目标检测和跟踪变得困难。其次，矿井的光照条件普遍较差，视频质量不稳定，部分视频较为模糊，尤其是在复杂环境下，图像质量的下降进一步加剧了目标检测与跟踪的难度。此外，矿井的背景信息往往较为单一，整体呈现暗灰色调，进一步加大了目标与背景的区分难度。在这种环境下，目标被遮挡的情况尤为严重，这对传统的基于外观特征、运动轨迹等信息的目标跟踪算法提出了严峻考验。

在国内外的矿井目标跟踪研究中，许多研究者尝试通过改进现有的目标检测和跟踪算法来解决遮挡问题。FENG 等^[1]基于双阶段检测器 Faster R-CNN 的框架，引入了注意力机制网络，提高了模型在目标被遮挡时的检测能力。程健等^[2]则在跟踪学习检测（TLD）框架基础上，提出了一种融合多特征的复杂场景动态目标跟踪算法，该算法使用融合多特征的 MeanShift 算法替代传统光流法，增强了在复杂场景下对目标形变的鲁棒性。陈太光等^[3]使用轻量化的全尺度特征学习 Re-ID 特征提取模型 OSNet 对 DeepSORT 进行优化，替换原有的 CNN 特征提取模块。陈伟等^[4]则提出了一种基于注意力机制的无监督矿井人员跟踪算法，结合了相关滤波和孪生网络在跟踪任务中的优势，成功构建了轻量化目标跟踪模型。然而当前开源的矿井遮挡人员目标跟踪数据集尚处于空白状态，国内外对矿井人员目标跟踪的研究主要依赖于小规模私有自建数据集。这些数据集的制作标准不统一，且大多数数据集的样本量都未超过 10000 张，难以为其他研究者提供足够的高质量数据支持。

为解决上述问题，本研究通过使用煤矿用本安型执法记录仪对矿井环境中的遮挡人员现场进行拍摄，整理构建了具有统一标准、大规模、高质量的矿井遮挡人员目标跟踪数据集 CUPOD (Coal mine Utility-area Personnel Occlusion-aware Dataset, CUPOD)^[5]。本数据集总共包含 14740 张图像，总数据量为 38.1GB，囊括了三种矿井典型人员出现遮挡的场景（出入口闸机（exit_gate）、井底硐室（well_chamber）和井下工作面（downhole_face））。其中 downhole_face 场景分辨率为 1280x720，exit_gate 和 well_chamber 场景分辨率为 1920x1080。该数据集的建立可有效缓解矿井遮挡人员目标跟踪数据资源不足的问题，具有较高应用价值，为相关研究提供了必要的技术支持。

1 数据采集和处理方法

为构建高质量、标准化的矿井遮挡人员目标跟踪数据集，本研究从数据采集、数据清洗到标注流程进行了系统设计与严格执行。首先，采用煤矿专用本安型执法记录仪在真实矿井环境中获取原始视频数据；其次，对原始数据进行清洗处理，剔除无效、低质量与涉隐数据；最后，基于 MOT 标注格式，借助 DarkLabel 软件对目标进行精准标注，确保数据具备良好的结构性与可用性。

1.1 数据采集方法

本数据于 2022 年 7 月至 2024 年 8 月在陕西省彬州市某煤矿采集完成。采集设备采用煤矿用本安型执法记录仪。数据采集设备的主要参数信息如表 1 所示。

为确保数据采集的稳定性和场景覆盖广度，所有视频均通过固定安装的执法记录仪进行拍摄，固定高度约为 1.5–2 米，镜头朝向覆盖通行区域或作业区域，确保完整捕捉目标人员的移动轨迹。采集过程中不存在人为干预或移动操作，避免了手持拍摄可能带来的抖动与视角变化。如图 1 所示，

通过煤矿用本安型执法记录仪进行多场景数据采集，模拟了矿工实际作业过程中视觉设备对目标的动态图像捕捉过程。

表 1 数据采集设备 DSJ-EX9 的主要参数信息

Table 1 Key specifications of the DSJ-EX9 intrinsically safe video recorder

参数名称	说明	参数名称	说明
设备型号	DSJ-EX9	视频格式	MPEG-4
设备类型	煤矿用本质安全型视音频记录仪	防爆认证编号	CQEx22.1119X
视频帧率	30 帧/秒 (FPS)	煤安认证编号	MFA170021
最大图像分辨率	400 万像素	制造商	廊坊市捷德电子科技有限公司



图 1 下井口出入口闸机（左）、井底硐室（中）和井下工作面（右）的图像数据展示图

Figure 1 Image samples captured at the underground entrance gate (left), the bottom chamber (middle), and the underground working face (right)

图 1 所示的三类典型场景分别分布于矿井不同功能区域：

下井口出入口闸机 (exit_gate) 位于下井口与主运输大巷之间，是井下人员进出作业系统的必经通道。该区域一般具备相对稳定的人工照明，光照条件较好，空间开阔，人流密集，目标遮挡频发，是跟踪任务中最具挑战性的场景之一。

井底硐室 (well_chamber) 指的是矿工在井下排队乘坐轨道车前往工作面前的候乘区域，一般位于井底车场附近。该区域空间较狭长，人员集中但活动范围小，遮挡情况适中，照明条件为人工固定灯带，亮度中等。该场景适用于评估遮挡、静态密集目标的短时间跟踪性能。

井下工作面 (downhole_face) 指的是煤矿作业现场最前端的生产区域，如采煤面或掘进头，照明条件较差，伴随大量扬尘、湿气和复杂背景结构。该区域图像质量普遍偏低，但目标遮挡相对较少，运动模糊也更显著。

数据采集的对象为矿井遮挡人员。数据采集主要环境背景及场景设计如表 2 所示。

表 2 数据集场景情况

Table 2 Dataset scene description

场景	视频总时长	场景说明
下井口出入口 闸机	23h25m21s	视频分辨率高；拍摄区域合适；遮挡情况较多；目标运动模糊程度较低；图像质量较优；低光照和局部曝光情况很少
井底硐室	1h1m20s	视频分辨率低；拍摄区域狭长；遮挡情况较少；目标运动模糊程度较高；图像质量差；低光照和局部曝光情况很少

场景	视频总时长	场景说明
井下工作面	23m29s	视频分辨率低；拍摄区域合适；遮挡情况很少；目标模糊程度较低；图像质量差；低光照和局部曝光情况严重

1.2 数据处理与标注

数据处理与标注是将采集到的原始视频数据进行系统性整理与预处理的过程，主要包括视频格式的统一转换、无效片段与污染数据的剔除，以及对感兴趣目标的人工标注操作^[6]。

1.2.1 数据清洗

本数据集共包括 11 个原始视频片段，总时长达 24 小时 50 分钟 10 秒。首先，利用人工筛选对原始视频序列进行初次筛选，剔除无效的视频序列。接着，对剩余的有效视频序列进行切帧处理，每个视频序列会根据需要被切分为 250 到 650 帧不等的多个小段，保存为独立的视频序列。然后，对这些切分后的视频序列进行数据清洗，数据清洗过程包含以下三个步骤：

（1）去除非目标数据：在数据采集过程中，部分视频序列未包含任何有效目标。这类数据对目标跟踪研究无实际意义，因此通过人工筛选将其剔除。

（2）剔除低质量数据：部分视频序列由于拍摄条件或设备限制，可能存在模糊、过度曝光或噪声较大等情况的图像帧。此类数据严重影响数据质量，因此被通过人工筛选方式剔除。

（3）隐私保护：隐私保护是数据集构建中的重要考量。为确保不侵犯相关公司的商业机密或引发任何产权问题，对所有涉及的公司标识及地点的信息进行了马赛克处理。

1.2.2 数据标注

本研究使用 DarkLabel 软件对清洗后的视频序列数据集进行标注。DarkLabel 软件支持 YOLO、MOT 以及 CreateML 等多种标注格式，能够广泛应用于计算机视觉和目标跟踪领域。YOLO (You Only Look Once) 格式数据集广泛应用于实时目标检测，具有高效且快速的检测能力^[7]；MOT (Multiple Object Tracking) 格式则用于多目标跟踪任务，能够帮助研究者对目标的运动轨迹进行标注^[8]。本研究基于 MOT 格式使用 DarkLabel 工具对图像数据集进行标注。

为确保标注数据的高质量，在数据标注过程中严格遵循以下 4 个标准：（1）标注时需贴合目标的实际边缘，以确保标注框精确。（2）在目标存在重叠或小范围遮挡的情况下，进行适当标注，确保目标的跟踪数据完整。（3）避免标注框过近图像边缘，以确保目标框的准确性和可追踪性。（4）剔除不包含有效目标的图像，确保每个标注帧都含有可用的目标数据。

2 数据样本描述

本数据集共包含 28 个数据段，数据段的命名从 downhole_face1-1 开始，依次包括 downhole_face1-2、exit_gate1-1 等，每个数据段的名称根据实际场景和视频序列进行命名。数据集分为 train 和 test 两部分，train 部分包含 21 个数据段，test 部分包含 7 个数据段。每个数据段主要包含视频序列和标注数据，下面分别对这两类数据的典型样本进行描述和分析。

2.1 图像数据样本

每个数据段的图像帧存储在 img1 文件夹中，图像格式为 PNG，分辨率为 1280×720 像素(px) 或 1920×1080 像素。每个图像文件的命名规则为 00000000.png、00000001.png 等，其中数字表示视频帧的编号。图像文件示例如图 2 所示。文件夹内的图像文件按照视频序列的帧顺序依次排列，图像大小在 1.18 MB 至 2.95 MB 之间。

名称	日期	类型	大小	标记
00000000.png	2025/1/19 1:30	PNG 图片文件	2,502 KB	
00000001.png	2025/1/19 1:30	PNG 图片文件	2,899 KB	
00000002.png	2025/1/19 1:30	PNG 图片文件	2,932 KB	
00000003.png	2025/1/19 1:30	PNG 图片文件	2,932 KB	
00000004.png	2025/1/19 1:30	PNG 图片文件	2,937 KB	
00000005.png	2025/1/19 1:30	PNG 图片文件	2,941 KB	
00000006.png	2025/1/19 1:30	PNG 图片文件	2,940 KB	
00000007.png	2025/1/19 1:30	PNG 图片文件	2,954 KB	
00000008.png	2025/1/19 1:30	PNG 图片文件	2,956 KB	
00000009.png	2025/1/19 1:30	PNG 图片文件	2,961 KB	
00000010.png	2025/1/19 1:30	PNG 图片文件	2,961 KB	
00000011.png	2025/1/19 1:30	PNG 图片文件	2,962 KB	
00000012.png	2025/1/19 1:30	PNG 图片文件	2,927 KB	
00000013.png	2025/1/19 1:30	PNG 图片文件	2,937 KB	
00000014.png	2025/1/19 1:30	PNG 图片文件	2,964 KB	
00000015.png	2025/1/19 1:30	PNG 图片文件	2,966 KB	
00000016.png	2025/1/19 1:30	PNG 图片文件	2,968 KB	
00000017.png	2025/1/19 1:30	PNG 图片文件	2,971 KB	
00000018.png	2025/1/19 1:30	PNG 图片文件	2,971 KB	

图 2 windows 文件系统下某数据段样本中图像数据详情

Figure 2 Image data details of a sample from a data segment in the windows file system

每个数据段名称、图像帧数、分辨率以及对应的目标实例数量详见表 3 所示。该表对 CUPOD 数据集中全部 28 个数据段进行了统计，可为后续模型训练与评估提供基础参考。

表 3 每个数据段名称及其相关信息表

Table 3 Names of each data segment and their associated information

序号	视频序列名称	图像帧数	分辨率(px)	目标实例数
train	downhole_face1-1	400	1280×720	1751
	downhole_face1-2	448	1280×720	2488
	downhole_face2	284	1280×720	1297
	exit_gate1-1	649	1920×1080	1955
	exit_gate1-2	649	1920×1080	2432
	exit_gate1-3	649	1920×1080	1471
	exit_gate2-1	639	1920×1080	5080
	exit_gate2-2	639	1920×1080	5818
	exit_gate2-3	639	1920×1080	4339
	exit_gate3-1	557	1920×1080	3340
	exit_gate3-2	557	1920×1080	5890

序号	视频序列名称	图像帧数	分辨率(px)	目标实例数
	exit_gate3-3	557	1920×1080	7547
	exit_gate3-4	557	1920×1080	9677
	exit_gate4-1	557	1920×1080	8990
	exit_gate4-2	557	1920×1080	9832
	exit_gate4-3	557	1920×1080	8979
	exit_gate4-4	557	1920×1080	10502
	well_chamber1-1	350	1920×1080	2724
	well_chamber1-2	406	1920×1080	3239
	well_chamber2-1	500	1920×1080	4288
	well_chamber2-2	513	1920×1080	3349
test	downhole_face3	428	1280×720	1482
	exit_gate1-4	649	1920×1080	2334
	exit_gate2-4	639	1920×1080	2999
	exit_gate3-5	558	1920×1080	9808
	exit_gate4-5	560	1920×1080	10298
	well_chamber3-1	342	1920×1080	3516
	well_chamber3-2	343	1920×1080	2062

2.2 标注数据样本

标注数据结果以*.txt 文件格式存储, 其中每个数据段根据实际场景和视频序列进行命名, 如 downhole_face1-1、downhole_face1-2、exit_gate1-1 等。每个数据段的标注数据按视频帧序列存储在单一的 txt 文件中。具体格式约定如表 4 所示。

标注文件的具体格式要求如下: (1) 每个图像序列的标注结果保存为一个文件。(2) 结果文件中所有英文标注均使用小写字母, 所有数字均使用阿拉伯数字。(3) 每个结果文件中的项都从 0 开始, 逐增编号, 同一帧中目标重复编号。(4) 同一帧内的多个目标应分行依次记录。(5) 同一数据行内不同字段之间以空格隔开。

表 4 数据标注文件格式说明

Table 4 Description of data annotation file format

	帧号	目标 ID	目标的 左上角 x 坐标	目标的 左上角 y 坐标	目标框 的宽度	目标框 的高度	置信度	目标类 别标志	目标的 可见性 比例	额外 信息
第一行	1	1	x 坐标	y 坐标	宽度	高度	-1	1	-1	-1
第二行	1	2	x 坐标	y 坐标	宽度	高度	-1	1	-1	-1
...	...									
第 i 行	m	n	x 坐标	y 坐标	宽度	高度	-1	1	-1	-1
...	...									
第 N 行	M	N	x 坐标	y 坐标	宽度	高度	-1	1	-1	-1

下面以数据集中 downhole_face1-2 视频序列下 gt 文件为例，对标注数据文件进行解释说明，如图 3 所示。

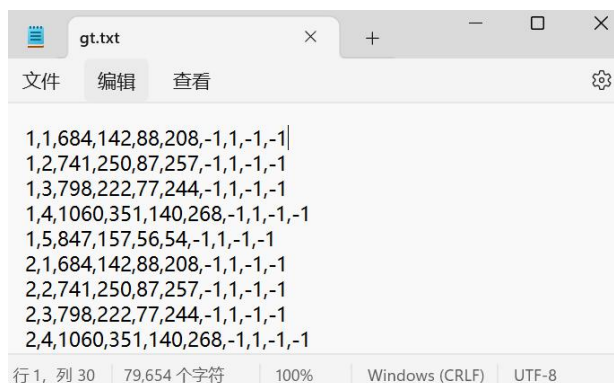


图 3 数据标注文件样本举例

Figure 3 Example of data annotation files

在第一行中，字段 1 表示该标注数据属于第 1 帧。目标 ID 为 1，表示这是第一个目标。接下来的字段“684”和“142”是该目标的左上角坐标 (x,y)，即目标所在位置的起始点。字段“88”和“208”是物体框的宽度和高度。字段“-1”表示置信度，在该数据集中-1 通常表示无置信度信息（即 GT 数据）。字段“1”表示该目标的类别 ID，这里可能代表该目标是“人员”或其他类别。字段“-1”代表该目标的可见性比例信息缺失或未提供。最后的字段“-1”表示可选扩展字段，但在此数据集中未使用。

在第二行中，字段“1”仍然表示该标注数据属于第 1 帧，目标 ID 为 2，表示这是第二个目标。字段“741”和“250”表示目标 2 的左上角坐标，字段“87”和“257”表示目标 2 的宽度和高度。字段“-1”依旧表示无置信度信息，字段“1”代表目标类别 ID，字段“-1”表示无可见性信息，最后的“-1”仍然是可选扩展字段（当前未使用）。

这种结构持续到每一帧的标注数据，数据按照帧号排列，所有目标的标注都在对应的帧内按序列出。每一帧的目标信息以行为单位进行存储，目标的编号、位置、尺寸、类别和可见性都会依次记录在该帧的数据行中。

2.3 数据集统计与分析

CUPOD 数据集共包含 137487 个标注的目标实例，其分布情况如图 4 所示。图 5 展示了数据集中目标尺寸的分布情况。本文以目标实例的像素面积作为尺度衡量指标，观察到数据集中目标的尺度范围从 6000 像素到 155000 像素不等。其中，大多数目标的面积超过 10000 像素，表明数据集在目标尺寸上具有较大的方差，能够满足研究人员对不同尺度目标检测与跟踪的多样性需求。

CUPOD 数据集数据采集时间跨度长，涉及多个场景，因此本数据集在多样性方面表现出色。本数据集涵盖了亮度变化、目标遮挡变化、目标模糊变化等多个维度的差异。如图 6 所示。

图 6(a)呈现了一些亮度变化的示例。由于数据集是从不同人员场景中捕获的，因此涉及多光源带来的亮度变化。图 6(b)展示了一些因拍摄环境和目标运动而引起目标模糊变化的示例。图 6(c)展示了不同程度目标遮挡的示例。

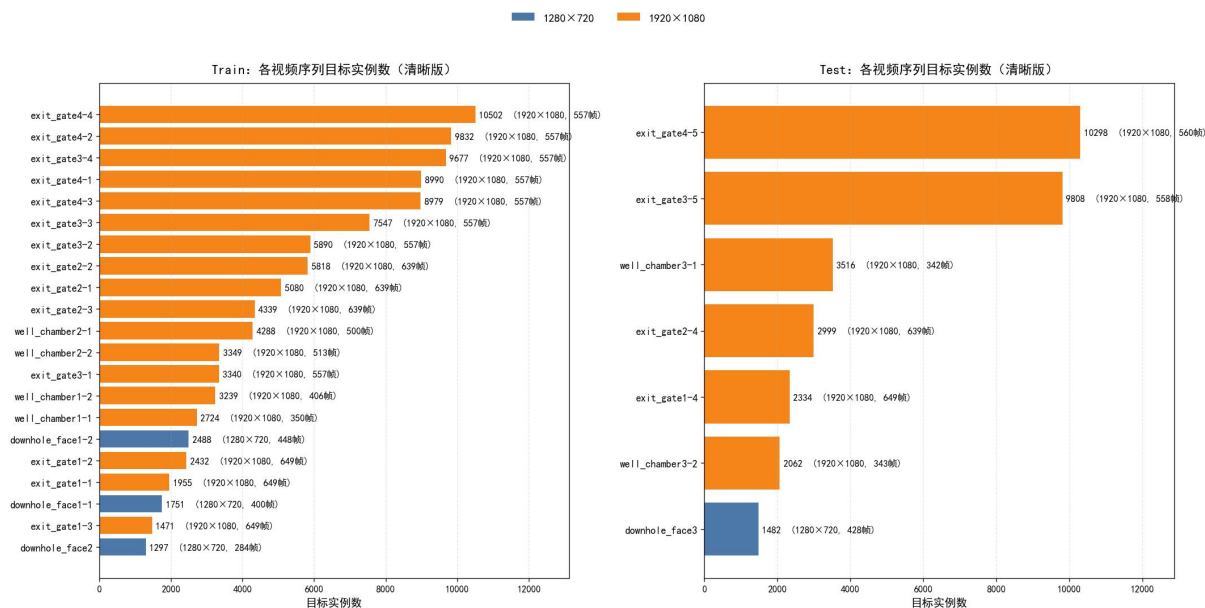


图4 完整数据集目标实例分布图

Figure 4 Distribution of target instances in the complete dataset

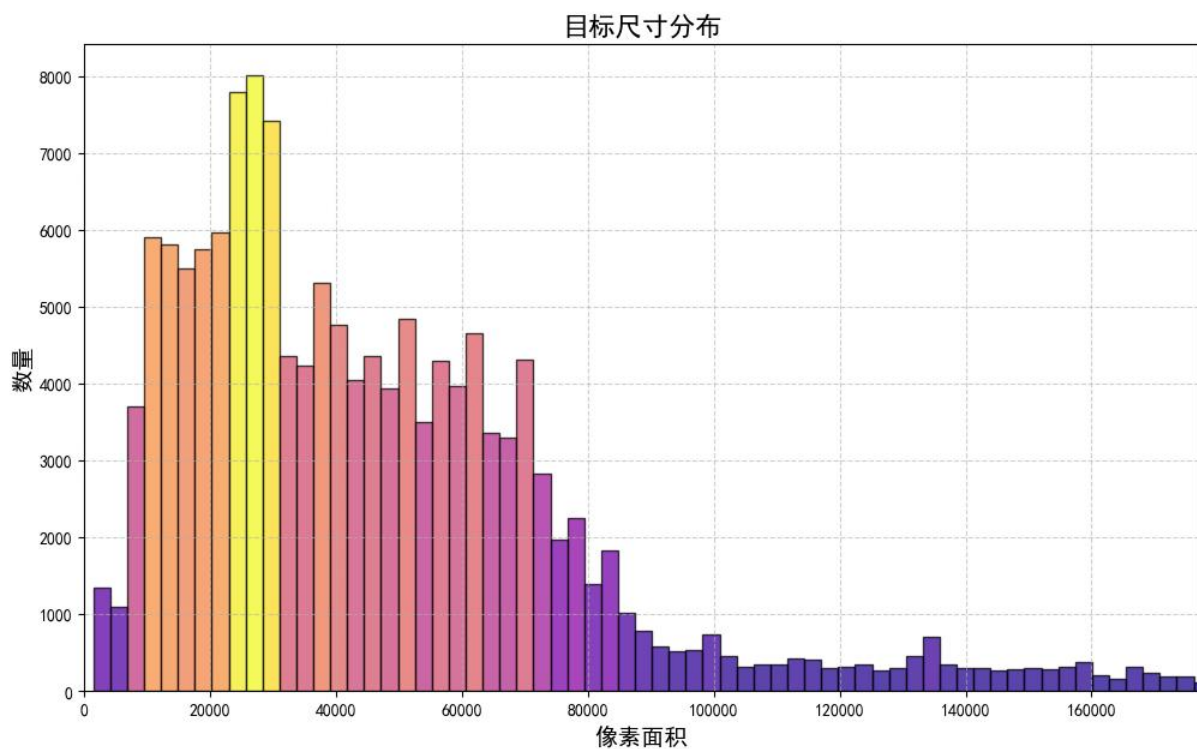


图5 目标尺寸分布图

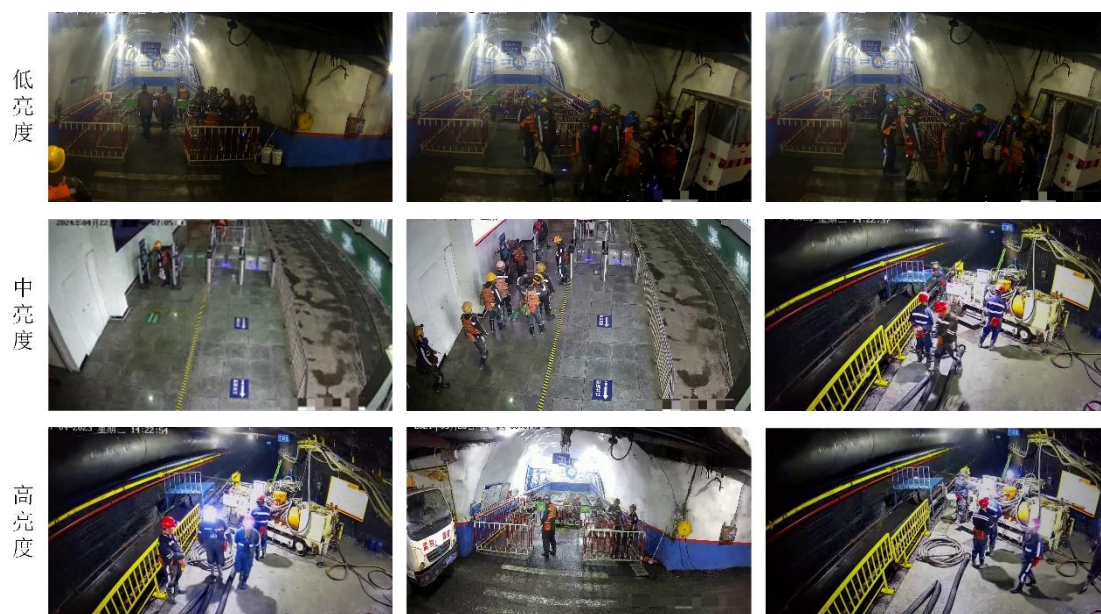
Figure 5 Distribution of target sizes

3 数据质量控制和评估

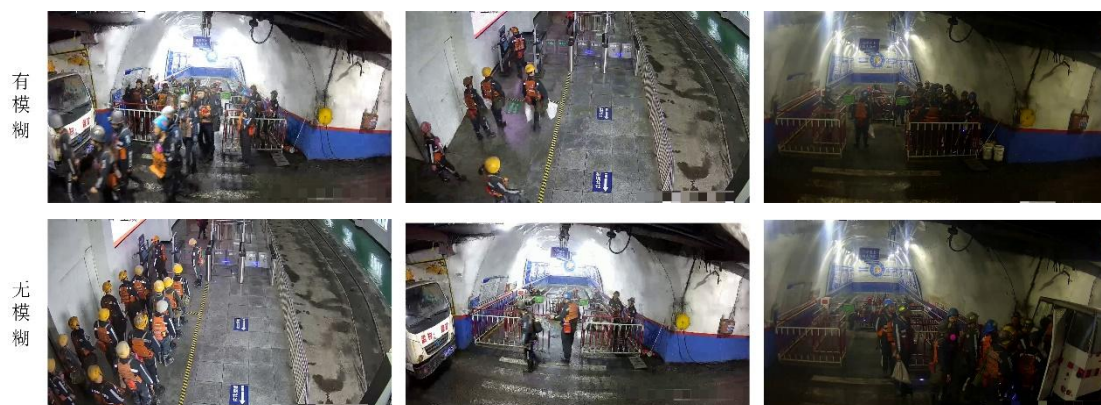
在本研究中，数据清洗和标注工作均由经过专业培训的人员严格按照标准操作流程执行，并有多名煤矿领域专家参与数据集的抽检与核查。为确保数据集的高质量，采取了以下措施：（1）通过

人工筛选和自动化工具剔除了无效或损坏的数据，确保数据集包含的是高质量的图像和标注信息。

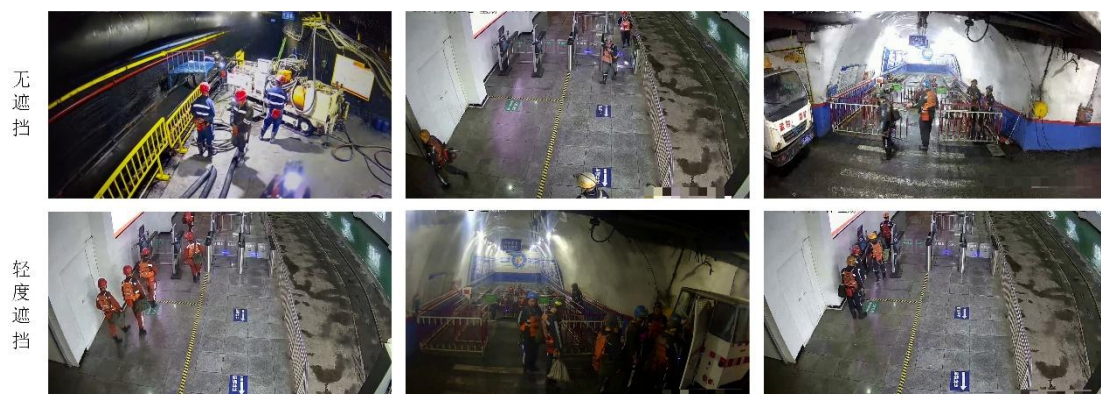
(2) 针对图像帧中的目标遮挡情况，采用了多次标注和交叉验证的方式，确保了标注数据的一致性和准确性。(3) 制定严谨的图像标注标准，以便保证数据集的标注质量和准确性。标注标准涵盖了目标边界框的准确定位、目标类别的准确识别以及各种特殊场景下的有效标注，以确保标注结果既能满足深度学习模型的训练需求，又能有效应对真实世界中的多样性和复杂性。(4) 多名煤矿领域专家参与数据集多轮抽检检查工作，确保数据集的可用性和准确性。

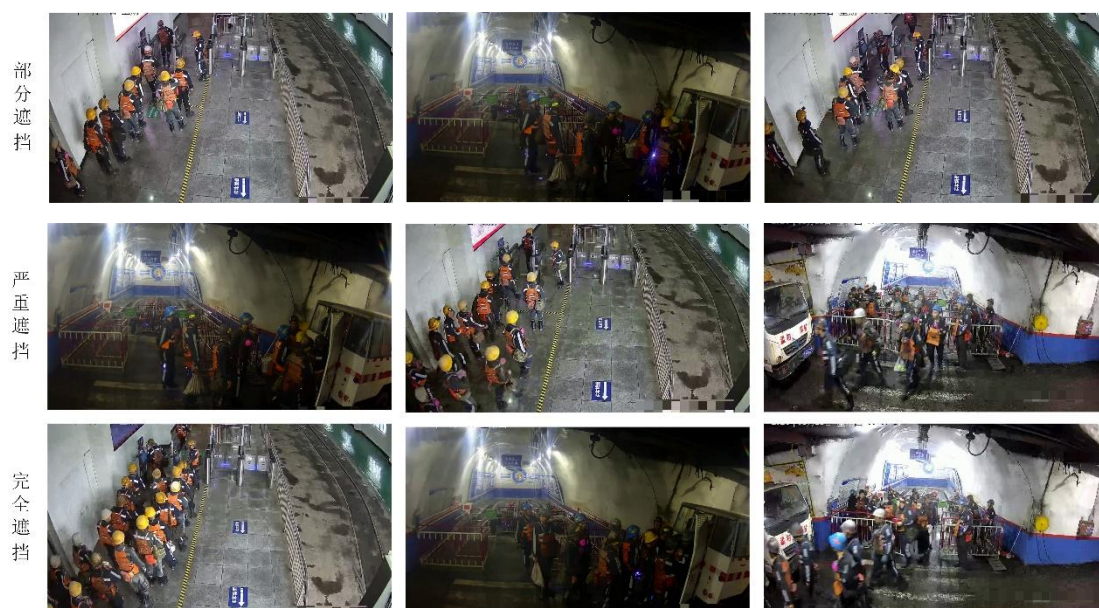


(a)不同人员场景下亮度差异示意图



(b) 不同人员场景下模糊差异示意图





(c) 不同人员场景下遮挡差异示意图

图 6 数据集多样性示例

Figure 6 Examples of dataset diversity

为了进一步验证数据集的有效性, 使用了 5 种先进的目标跟踪算法对 CUDSD 数据集进行跟踪评估实验, 具体包括: ImpAssoc、OC-SORT^[9]、ByteTrack^[10]、StrongSORT^[11]和 BoT-SORT^[12]。为系统验证 CUDSD 数据集在复杂场景下的跟踪任务适配性, 本研究基于 BoxMOT 开源框架构建了标准化评测体系。实验采用控制变量法设计, 通过固定检测与重识别 (ReID) 基模型, 对主流跟踪算法进行横向对比分析。对于检测基准模型, 本研究选用 yolov8m.pt、yolov8x.pt、yolox_s.pt、yolox_l.pt 预训练模型; 对于 ReID 模型, 本研究采用 osnet_x0_25_msmt17.pt 和 osnet_x1_0_msmt17.pt、mobilenetv2_x1_4_market1501.pt 以及 clip_market1501.pt, 验证特征提取对跟踪性能的敏感性。本研究的实验环境配置如下: CPU 为 Intel(R) Xeon(R) Gold 6330 CPU@2.00GHz, GPU 为 RTX 3090, 显存为 24 GB, 操作系统为 Ubuntu 20.04, CUDA 版本为 11.3, Pytorch 版本为 1.10.0, Python 版本为 3.8。

目标跟踪任务的性能评估指标包括 HOTA (Higher Order Tracking Accuracy)、MOTA (Multiple Object Tracking Accuracy)、IDF1 (ID F1-Score)、MOTP (Multiple Object Tracking Precision)、MODA (Multiple Object Detection Accuracy)、DetA (Detection Accuracy) 和 AssA (Association Accuracy)。其中, HOTA 综合衡量检测和关联精度, 是一种更高阶的目标跟踪评价标准。MOTA 主要衡量跟踪系统的整体准确性, 考虑了漏检、误检以及身份切换 (ID switch)。IDF1 衡量目标身份的一致性, 反映了跟踪过程中 ID 分配的正确性。MOTP 评估检测框的定位精度, 通常通过计算预测框与真实框的平均 IoU 来衡量。MODA 主要关注目标检测的准确性, 与 MOTA 类似, 但不考虑身份切换因素。DetA 衡量目标检测的精度, 即正确检测到目标的比例。AssA 衡量目标关联的准确性, 即正确跟踪目标身份的比例。

HOTA 的计算方式如下:

$$HOTA = \sqrt{DetA \times AssA} \quad (1)$$

其中, $DetA$ 和 $AssA$ 分别表示检测精度和关联精度。

MOTA 和 MOTP 计算方式如下：

$$MOTA = 1 - \frac{FN + FP + ID_{sw}}{GT} \quad (2)$$

$$MOTP = \frac{\sum IoU}{\sum D} \quad (3)$$

其中， FN 为未检测目标的数量， FP 为误检目标的数量， ID_{sw} 为身份切换次数， GT 为真实目标数， IoU 表示检测框与真实框的交并比， D 为检测到的目标数量。

IDF1 计算方式如下：

$$IDF1 = \frac{2 \times IDTP}{2 \times IDTP + IDFP + IDFN} \quad (4)$$

其中， $IDTP$ 、 $IDFP$ 、 $IDFN$ 分别表示身份正确跟踪、错误跟踪以及未跟踪的数量。

MODA 计算方式如下：

$$MODA = 1 - \frac{FN + FP}{GT} \quad (5)$$

这些指标可以用于衡量目标跟踪算法的整体性能，评估其检测能力、关联能力以及精度。

CUPOD 数据集的基准评估结果如表 5 所示。在本次目标跟踪实验中，评估了不同跟踪算法在不同检测模型和重识别（ReID）模型组合下的性能，并采用 HOTA、MOTA、IDF1、MOTP、MODA、DetA、AssA 等指标进行对比分析。

表 5 CUPOD 数据集的基准评估结果

Table 5 Benchmark evaluation results on the CUPOD dataset

算法	检测模型	重识别模型	HOTA	MOTA	IDF1	MOTP	MODA	DetA	AssA
ImpAssoc	yolov8x	osnet_x1_0_msmt17	48.54	57.594	55.714	83.333	58.699	51.175	46.691
OC-SORT	yolov8x	osnet_x1_0_msmt17	48.484	49.282	59.397	84.694	49.387	42.412	55.827
ByteTrack	yolov8x	osnet_x1_0_msmt17	52.821	57.093	66.371	83.005	57.271	48.863	57.661
StrongSORT	yolov8x	osnet_x1_0_msmt17	52.056	40.244	60.622	81.762	41.325	49.352	56.009
BoT-SORT	yolov8x	osnet_x1_0_msmt17	54.071	55.35	67.183	83.332	55.447	47.601	61.868
ImpAssoc	yolov8m	osnet_x1_0_msmt17	43.762	53.373	50.067	82.751	54.576	47.17	41.424
OC-SORT	yolov8m	osnet_x1_0_msmt17	46.005	44.303	56.73	84.628	44.408	37.739	56.493
ByteTrack	yolov8m	osnet_x1_0_msmt17	49.178	53.333	61.338	82.475	53.543	44.99	54.457
StrongSORT	yolov8m	osnet_x1_0_msmt17	50.391	41.567	59.366	80.473	42.818	48.724	53.375
BoT-SORT	yolov8m	osnet_x1_0_msmt17	50.742	52.308	62.566	83.185	52.453	44.453	58.44
ImpAssoc	yolov8x	osnet_x0_25_msmt17	48.54	57.594	55.714	83.333	58.699	51.175	46.691
OC-SORT	yolov8x	osnet_x0_25_msmt17	48.484	49.282	59.397	84.694	49.387	42.412	55.827
ByteTrack	yolov8x	osnet_x0_25_msmt17	52.821	57.093	66.371	83.005	57.271	48.863	57.661

算法	检测模型	重识别模型	HOTA	MOTA	IDF1	MOTP	MODA	DetA	AssA
StrongSORT	yolov8x	osnet_x0_25_msmt17	51.373	40.284	59.413	81.727	41.406	49.343	54.644
BoT-SORT	yolov8x	osnet_x0_25_msmt17	54.071	55.35	67.183	83.332	55.447	47.601	61.868
ImpAssoc	yolov8m	osnet_x0_25_msmt17	43.71	53.349	49.94	82.752	54.567	47.203	41.304
OC-SORT	yolov8m	osnet_x0_25_msmt17	46.005	44.303	56.73	84.628	44.408	37.739	56.493
ByteTrack	yolov8m	osnet_x0_25_msmt17	49.178	53.333	61.338	82.475	53.543	44.99	54.457
StrongSORT	yolov8m	osnet_x0_25_msmt17	50.604	41.664	59.241	80.509	42.915	49.159	53.358
BoT-SORT	yolov8m	osnet_x0_25_msmt17	50.742	52.308	62.566	83.185	52.453	44.453	58.44
ImpAssoc	yolox_s	osnet_x1_0_msmt17	47.022	45.846	45.579	57.699	81.44	50.981	43.671
OC-SORT	yolox_s	osnet_x1_0_msmt17	48.752	48.317	43.441	53.115	82.497	44.555	53.471
ByteTrack	yolox_s	osnet_x1_0_msmt17	53.618	53.186	48.023	60.144	81.289	50.374	57.305
StrongSORT	yolox_s	osnet_x1_0_msmt17	47.16	45.518	20.568	33.231	81.032	42.338	52.882
BoT-SORT	yolox_s	osnet_x1_0_msmt17	52.614	51.981	48.086	59.902	81.532	50.396	55.157
ImpAssoc	yolox_s	osnet_x0_25_msmt17	47.026	45.881	45.617	57.739	81.438	50.945	43.704
OC-SORT	yolox_s	osnet_x0_25_msmt17	48.752	48.317	43.441	53.115	82.497	44.555	53.471
ByteTrack	yolox_s	osnet_x0_25_msmt17	53.618	53.186	48.023	60.144	81.289	50.374	57.305
StrongSORT	yolox_s	osnet_x0_25_msmt17	48.008	46.295	20.727	33.393	81.047	42.555	54.501
BoT-SORT	yolox_s	osnet_x0_25_msmt17	52.614	51.981	48.086	59.902	81.532	50.396	55.157
ImpAssoc	yolox_l	mobilenetv2_x1_4_market1501	50.039	48.449	45.666	56.609	83.662	52.112	48.481
OC-SORT	yolox_l	mobilenetv2_x1_4_market1501	55.043	53.995	48.22	57.892	84.295	50.958	59.765
ByteTrack	yolox_l	mobilenetv2_x1_4_market1501	53.986	52.894	47.857	58.974	83.392	52.48	55.962
StrongSORT	yolox_l	mobilenetv2_x1_4_market1501	50.582	48.469	37.539	48.677	83.457	49.397	52.31
BoT-SORT	yolox_l	mobilenetv2_x1_4_market1501	54.869	54.034	49.187	60.055	83.666	52.898	57.251
ImpAssoc	yolox_l	clip_market1501	50.29	48.705	45.75	56.698	83.659	52.169	48.904
OC-SORT	yolox_l	clip_market1501	55.043	53.995	48.22	57.892	84.295	50.958	59.765
ByteTrack	yolox_l	clip_market1501	53.986	52.894	47.857	58.974	83.392	52.48	55.962
StrongSORT	yolox_l	clip_market1501	52.432	50.188	37.952	49.096	83.469	49.559	56.027
BoT-SORT	yolox_l	clip_market1501	54.857	54.019	49.195	60.063	83.666	52.904	57.22

在 HOTA 指标上, OC-SORT (YOLOX-L+clip_market1501) 取得了 55.043 的最佳成绩, 表明该方法在检测与关联精度的平衡性方面表现最优。而在身份一致性 (IDF1) 方面, BoT-SORT (YOLOv8x+osnet_x1_0_msmt17) 以 67.183% 领先, 说明其在保持目标身份连续性上具有优势。

在整体跟踪准确率 (MOTA) 方面, OC-SORT (YOLOX-L+clip_market1501) 和 OC-SORT (YOLOX-L+mobilenetv2_x1_4_market1501) 均达到了 53.995%, 表现出良好的跟踪准确性。而在检测精度 (MOTP) 方面, OC-SORT (YOLOv8x+osnet_x1_0_msmt17) 以 84.694% 领先, 显示其在目标检测框精度上的优势。

综合来看, OC-SORT (YOLOX-L+clip_market1501) 在 HOTA (55.043%) 和 MOTA (53.995%)

两个核心指标上均取得最高分，说明该模型在检测与关联方面表现最佳，适用于高精度跟踪任务。而 BoT-SORT (YOLOv8x+osnet_x1_0_msmt17) 虽然 HOTA 略低 (54.071%)，但在 IDF1 (67.183%) 上领先，适用于身份一致性要求较高的任务。

4 数据价值

本研究通过拍摄矿井遮挡人员目标，构建了一个真实且多样化的矿井目标跟踪数据集，旨在为该领域的相关研究提供可靠的数据支持。目前，国内矿井目标跟踪数据集普遍规模较小且标注不统一，样本量通常不足 10000 帧，且数据集的制作缺乏统一的标准和规范。与现有数据集相比，本数据集在制作过程中严格遵循高质量标准，并经过多次专家检查和验证，确保了数据集的高质量和高可用性。此外，本数据集包含 28 个数据段，图像帧总数超过 15000 帧，目标实例数量达到 137487 个，标注数据细致，涵盖了目标跟踪中涉及的不同遮挡场景。相比于其他矿井目标跟踪数据集，本数据集提供了更多的样本量和更高质量的标注，能够为深度学习模型的训练提供高质量、全面的数据支持，并且可以为矿井人员目标跟踪领域的研究提供更具广泛适用性的数据集。

5 数据使用方法和建议

本数据集已按照目标跟踪任务的需求进行预先划分，分为训练集和测试集两部分，训练集包含 21 个数据段，测试集包含 7 个数据段。每个数据段中，数据按帧进行标注，包含目标的检测结果和跟踪标注数据。数据格式为 txt 文件，已按照标准的目标跟踪格式 (MOT 格式) 进行处理，支持与多种深度学习框架兼容。

用户可以根据自身任务需求选择训练集或测试集进行模型训练和验证。本数据集特别适用于目标跟踪模型的训练，可以在深度学习框架中直接使用。用户也可以通过脚本对数据进行进一步处理，如按特定需求划分数据集或进行数据增强。

此外，为提高数据集在目标检测任务中的适用性，本研究提供了从 MOT 格式向 YOLO 格式的转换脚本。该脚本基于每帧图像的标注信息，提取目标的左上角坐标及目标框的宽高，计算出边界框的中心点坐标，并将其转换为相对于图像尺寸归一化后的 YOLO 格式 (即 class_id、x_center、y_center、width、height)，所有坐标信息均为小数形式。转换结果将以与图像同名的 txt 文件存储，便于直接用于 YOLO 系列检测模型的训练。

作者分工职责

高丽铄 (2000—)，女，河北省邢台市人，硕士，助教，研究方向为计算机视觉。主要承担工作：数据采集、数据标注、数据清洗、数据汇总整理、论文撰写。

路洋 (2000—)，男，安徽省合肥市人，硕士研究生，研究方向为计算机视觉。主要承担工作：数据采集、数据标注、论文撰写。

董立红 (1968—)，女，陕西省西安市人，博士，教授，研究方向为智能信息处理技术。主要承担工作：总体方案设计与组织实施。

叶鸥 (1984—)，男，陕西省西安市人，博士，讲师，研究方向为人工智能。主要承担工作：提出数据采集规范、论文修改。

韩冷雨（1999—），男，安徽省合肥市人，硕士研究生，研究方向为计算机视觉。主要承担工作：数据采集、数据标注。

郝晨阳（2001—）男，陕西省渭南市人，硕士研究生，研究方向为计算机视觉。主要承担工作：数据采集、数据标注。

张立智（1999—）男，山西省太原市人，硕士研究生，研究方向为抗干扰，无源定位。主要承担工作：数据采集、数据标注。

参考文献

- [1] 冯磊, 李斌, 何勇. 基于注意力机制的煤矿工作面遮挡行人检测[J]. 价值工程, 2022, 41(14): 148 - 150. DOI:10.3969/j.issn.1006-4311.2022.14.047. [FENG L, LI B, HE Y. Pedestrian detection of coal mine working face occlusion based on attention mechanism[J]. Value Engineering, 2022, 41(14): 148 - 150. DOI:10.3969/j.issn.1006-4311.2022.14.047.]
- [2] 程健, 陈亮, 王凯, 等. 一种多特征融合的复杂场景动态目标跟踪算法[J]. 中国矿业大学学报, 2021, 50(5): 1002 - 1010. DOI:10.13247/j.cnki.jcumt.001333. [CHENG J, CHEN L, WANG K, et al. Multi-feature fusion dynamic target tracking algorithm for complex scenes[J]. Journal of China University of Mining & Technology, 2021, 50(5): 1002 - 1010. DOI:10.13247/j.cnki.jcumt.001333.]
- [3] 陈太光, 鲍新平, 汪涛, 等. 基于深度学习的矿井巷道人员计数技术[J]. 煤矿安全, 2023, 54(2): 234 - 238. DOI:10.13347/j.cnki.mkaq.2023.02.035. [CHEN T G, BAO X P, WANG T, et al. Mine roadway personnel counting technology based on deep learning[J]. Safety in Coal Mines, 2023, 54(2): 234 - 238. DOI:10.13347/j.cnki.mkaq.2023.02.035.]
- [4] 陈伟, 任鹏, 田子建, 等. 基于注意力机制的无监督矿井人员跟踪[J]. 煤炭学报, 2021, 46(S1): 601 - 608. DOI:10.13225/j.cnki.jccs.2020.1348. [CHEN W, REN P, TIAN Z J, et al. Unsupervised mine personnel tracking based on attention mechanism[J]. Journal of China Coal Society, 2021, 46(S1): 601 - 608. DOI:10.13225/j.cnki.jccs.2020.1348.]
- [5] 周蔚, 董立红, 叶鸥, 等. 煤矿井下钻场目标检测数据集[J/OL]. 中国科学数据, 2024, 9(2): 300 - 312. DOI:10.11922/11-6035.csd.2024.0006.zh. [ZHOU W, DONG L H, YE O, et al. A dataset of drilling site object detection in underground coal mines[J/OL]. China Scientific Data, 2024, 9(2): 300 - 312. DOI:10.11922/11-6035.csd.2024.0006.zh.]
- [6] 回丙伟, 宋志勇, 范红旗, 等. 地/空背景下红外图像弱小飞机目标检测跟踪数据集[J/OL]. 中国科学数据, 2020, 5(3): 286 - 297. DOI:10.11922/csdata.2019.0074.zh. [HUI B W, SONG Z Y, FAN H Q, et al. A dataset for infrared detection and tracking of dim-small aircraft targets under ground/air background[J/OL]. China Scientific Data, 2020, 5(3): 286 - 297. DOI:10.11922/csdata.2019.0074.zh.]
- [7] 董红召, 林少轩, 余翊妮. 交通目标 YOLO 检测技术的研究进展[J]. 浙江大学学报(工学版), 2025, 59(2): 249 - 260. DOI:10.3785/j.issn.1008-973X.2025.02.003. [DONG H Z, LIN S X, SHE Y N. Research progress of YOLO detection technology for traffic object[J]. Journal of Zhejiang University (Engineering Science), 2025, 59(2): 249 - 260. DOI:10.3785/j.issn.1008-973X.2025.02.003.]
- [8] 闫晨阳, 刘宏哲, 徐成, 等. 面向智能驾驶的行人多目标跟踪算法研究[J]. 计算机工程与应用, 2023, 59(15): 206 - 213. DOI:10.3778/j.issn.1002-8331.2204-0105. [YAN C Y, LIU H Z, XU C, et al.

Research on pedestrian multi-object tracking algorithm for intelligent driving[J]. Computer Engineering and Applications, 2023, 59(15): 206 – 213. DOI:10.3778/j.issn.1002-8331.2204-0105.]

[9] MAGGIOLINO G, AHMAD A, CAO J K, et al. Deep OC-sort: multi-pedestrian tracking by adaptive re-identification[C]//2023 IEEE International Conference on Image Processing (ICIP). Kuala Lumpur, Malaysia. IEEE, 2023: 3025 – 3029. DOI:10.1109/ICIP49359.2023.10222576.

[10] ZHANG Y F, SUN P Z, JIANG Y, et al. ByteTrack: multi-object tracking by associating every detection box[C]//Computer Vision – ECCV 2022. ACM, 2022: 1 – 21. DOI:10.1007/978-3-031-20047-2_1.

[11] DU Y H, ZHAO Z C, SONG Y, et al. StrongSORT: make DeepSORT great again[J]. IEEE Transactions on Multimedia, 2023, 25: 8725 – 8737. DOI:10.1109/TMM.2023.3240881.

[12] AHARON N, ORFAIG R, BOBROVSKY B Z. BoT-SORT: robust associations multi-pedestrian tracking[EB/OL]. 2022: arXiv: 2206.14651. <https://arxiv.org/abs/2206.14651>.

论文引用格式

高丽铄, 路洋, 董立红, 等. 矿井遮挡人员跟踪数据集[J/OL]. 中国科学数据, 2026, 11(1). (2025-12-29). DOI: 10.11922/11-6035.csd.2025.0028.zh.

数据引用格式

高丽铄, 路洋, 董立红, 等. 矿井遮挡人员跟踪数据集[DS/OL]. V3. Science Data Bank, 2025. (2025-12-18). DOI:10.57760/sciencedb.21999.

A dataset of occlusion-aware personnel tracking in coal mine utility areas

GAO Lishuo¹, LU Yang^{2*}, DONG Lihong², YE Ou², HAN Lengyu², ZHANG Lizhi², XI Chenyang²

1. Xi'an International University, Xi'an 710077, P.R. China

2. Xi'an University of Science and Technology, Xi'an 710054, P.R. China

*Email: luyang20000520@163.com

Abstract: The complex and unique operational environment of underground mines presents significant challenges for personnel target detection and tracking. Workers typically wear standardized work uniforms, resulting in highly similar visual appearances. In addition, insufficient illumination, motion blur, and monotonous gray backgrounds significantly reduce the accuracy of visual recognition. In typical underground mining scenarios, high personnel density and frequent occlusions further increase the difficulty of tracking algorithms. In deep learning model training, the quality of datasets often plays a decisive role determining model performance. However, there is currently a lack of publicly available and standardized image datasets for personnel tracking in mining environments, both domestically and internationally. Most existing studies rely on small-scale, non-standardized private datasets, which are

insufficient for comprehensive algorithm benchmarking and large-scale application. To address this gap, this study utilizes intrinsically safe law enforcement recorders designed for coal mines to capture real-world occlusion scenarios in underground mining environments. Based on the collected data, a large-scale, standardized, and high-quality occlusion-aware personnel target tracking dataset for mines is constructed. The dataset comprises a total of 14,740 images with an overall size of 38.1 GB, covering three typical occlusion-prone scenarios in mines: the exit gate at the shaft entrance, the well chamber at the bottom of the shaft, and the downhole working face. The establishment of this dataset effectively alleviates the scarcity of occluded personnel tracking resources in mining environments and holds high practical value. It provides essential data support for related research and facilitates the development and evaluation of advanced tracking algorithms in challenging real-world mining scenarios.

Keywords: coal mine; object tracking; occluded personnel; deep learning; dataset

Dataset Profile

Title	A dataset of occlusion-aware personnel tracking in coal mine utility areas
Data corresponding author	LU Yang (luyang20000520@163.com)
Data authors	GAO Lishuo, LU Yang, DONG Lihong, YE Ou, HAN Lengyu, ZHANG Lizhi, XI Chenyang
Time range	2022–2024
Geographical scope	Binzhou City, Shaanxi Province
Data volume	38.1 GB
Data format	*.png, *.txt
Data service system	https://doi.org/10.57760/sciencedb.21999
Source of funding	National Natural Science Foundation of China (62303375)
Dataset composition	This dataset consists of 28 data segments, each containing 3 parts: (1) The det.txt file stores the detection results generated by an object detection algorithm. The fields include the image frame number, motion trajectory ID, target position coordinates, bounding box width and height, and the range flag value for the corresponding data segment. (2) The gt.txt file contains the tracking annotation data for each frame of the image. The annotation includes the image frame number, motion trajectory ID, target position coordinates, bounding box width and height, occlusion status, class ID, and visibility ratio for each time point at intervals of 33.3 ms in the corresponding data segment. (3) The img1 folder contains the image data for each frame in the video sequence.