www.nature.com/scientificdata

scientific data

Check for updates

OPEN

DATA DESCRIPTOR

An open paradigm dataset for intelligent monitoring of underground drilling operations in coal mines

Pengzhen Zhao[1,2], Xichao Wang[1,2] ✉, Shuainan Yu[1,2], Xiangqing Dong[1,2], Baojiang Li[1,2], Haiyan Wang[1,2] & Guochu Chen[1,2]

The underground drilling environment in coal mines is critical and prone to accidents, with common accident types including rib spalling, roof falling, and others. High-quality datasets are essential for developing and validating artificial intelligence (AI) algorithms in coal mine safety monitoring and automation field. Currently, there is no comprehensive benchmark dataset for coal mine industrial scenarios, limiting the research progress of AI algorithms in this industry. For the first time, this study constructed a benchmark dataset (DsDPM 66) specifically for underground coal mine drilling operations, containing 105,096 images obtained from surveillance videos of multiple drilling operation scenes. The dataset has been manually annotated to support computer vision tasks such as object detection and pose estimation. In addition, this study conducted extensive benchmarking experiments on this dataset, applying various advanced AI algorithms including but not limited to YOLOv8 and DETR. The results indicate the proposed dataset highlights areas for improvement in algorithmic models and fills the data gap in the coal mining, providing valuable resources for developing coal mine safety monitoring.

Background & Summary

In the coming decades, coal resources will continue to be a major global energy source, and the safety and efficiency of coal mining will be the core focus of mining management[1,2]. In recent years, with the development of artificial intelligence technology, the application in coal mining has shown great potential by assisting or replacing humans in completing high-risk tasks, thereby enhancing safety monitoring and operational efficiency[3]. Intelligent mining systems are essential for ensuring safe and efficient underground production while addressing the high risks of underground operations. However, the actual application effect of artificial intelligence is limited by the availability and quality of data in specific scenarios and still relies on human involvement when handling complex tasks.

Among the various work scenarios in coal mining, the drilling operation in the heading face represents a high-risk environment prone to accidents such as rib spalling, roof falls, and personnel intrusion into equipment[4]. As illustrated in the research scenario analysis in Fig. 1, rib spalling refers to the detachment of coal or rock from the sidewalls of mining roadways, while roof falling involves the sudden collapse of overhead strata in underground spaces. These incidents pose severe safety risks, including worker injuries, equipment damage, and operational disruptions. Traditional monitoring methods primarily rely on manual observation and rule-based anomaly detection. However, these approaches are prone to human error and have limited adaptability to complex and dynamic underground environments. Manual monitoring requires continuous attention from professional staff, which can lead to fatigue-induced oversight and inconsistencies in anomaly detection. Additionally, existing video surveillance systems lack intelligent analysis capabilities and can only provide passive recordings without real-time intervention. Despite the urgent demand for automated safety monitoring, current AI-based solutions are hindered by insufficient domain-specific training data, which limits their generalization ability in real-world applications.

[1]School of Electrical Engineering, Shanghai DianJi University, Shanghai, 201306, China. [2]Intelligent Decision and Control Technology Institute, Shanghai DianJi University, Shanghai, 201306, China. ✉e-mail: wangxc@sdju.edu.cn

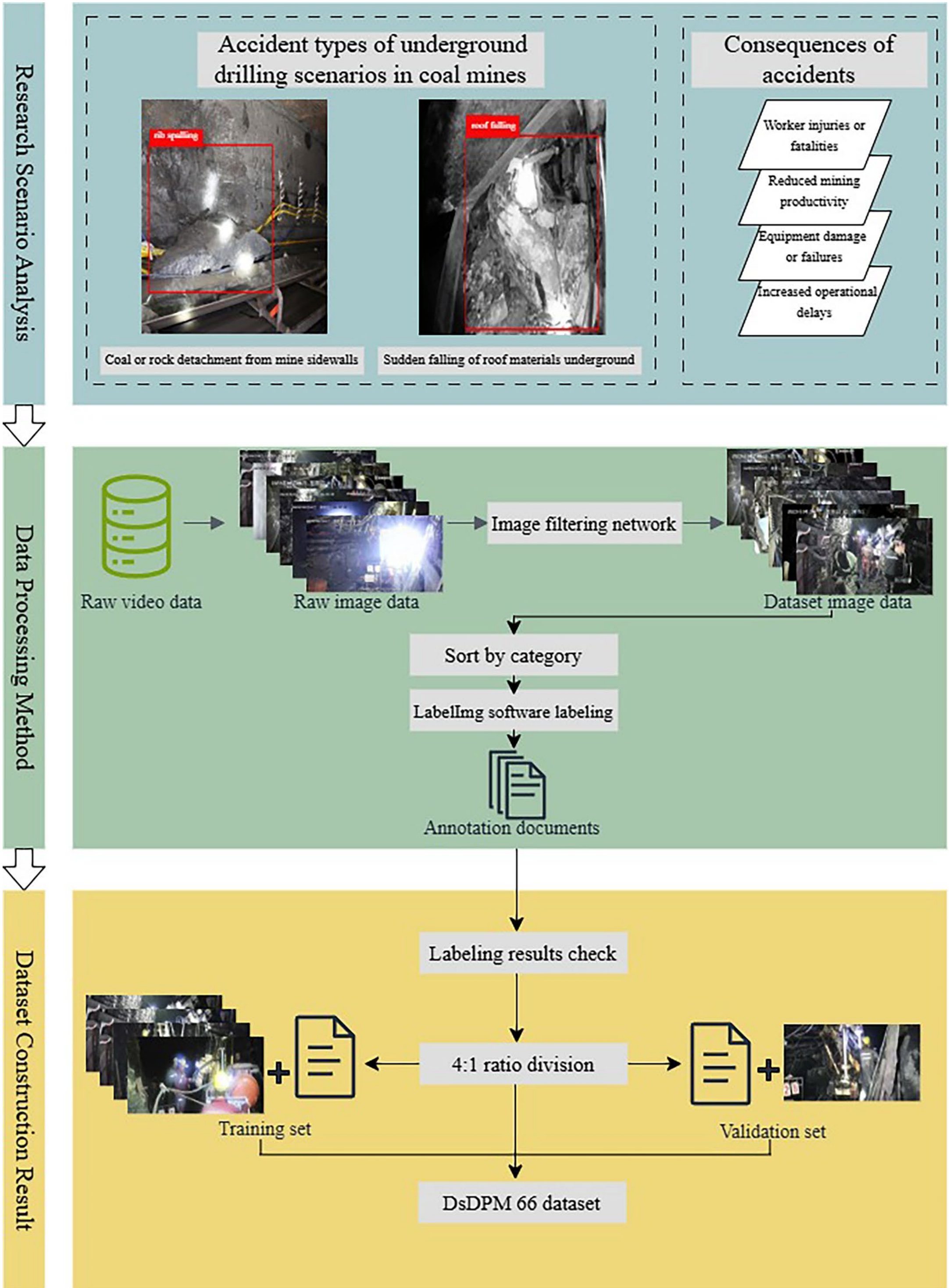


Fig. 1 Workflow of accident analysis, data processing, and the construction of the DsDPM 66 dataset for underground drilling operations in coal mines.

In drilling operations at the heading face, both equipment and personnel play a critical role in ensuring safety and maintaining production efficiency. Among the key components, drill pipes and drill rigs constitute the core of underground drilling systems, directly influencing operational stability and effectiveness. Drill pipes serve as essential structural elements, providing the necessary depth and mechanical support during drilling. By connecting multiple drill pipes in series, operators can extend the drilling reach to facilitate geological exploration or coal seam extraction. In this system, the drill rig itself does not engage in direct rock-breaking; instead, it transmits rotational force to the drill pipes, enabling controlled penetration of underground rock layers. Given their central function, accurately identifying and tracking drill pipes and drill rigs in real time is crucial for assessing their positions and operational statuses, thus preventing potential hazards. The interaction between miners and drill pipes primarily involves pipe dismantling and handling, tasks that require precise coordination to maintain operational continuity. During dismantling, miners must manually detach and maneuver individual drill pipes from the rig, ensuring seamless workflow while minimizing safety risks. The process carries inherent dangers, including the risk of falling pipes, which may cause severe injuries. Therefore, deploying AI-driven monitoring systems to analyze worker-pipe interactions is essential for enhancing safety management.

Miners are directly responsible for executing drilling tasks, and their movements, behaviors, and interactions with the surrounding environment serve as key indicators of work progress and safety compliance. To improve workplace safety, intelligent monitoring systems must be capable of real-time miner detection, determining whether workers remain within designated safe zones. Automated systems should promptly identify unauthorized entries into hazardous areas, issue auditory alerts, and, when necessary, trigger emergency shutdowns of drilling equipment. Moreover, personal protective equipment, such as safety helmets and compressed oxygen self-rescuers, is fundamental to miner safety in underground environments. The mining safety helmet provides critical protection against head injuries from falling debris, while the compressed oxygen self-rescuer serves as a vital life-support device in oxygen-deficient or toxic gas conditions. Given the significance of personal protective equipment compliance, AI-based identification systems can play a pivotal role in monitoring the proper use of protective gear. By promptly detecting improper helmet usage or the absence of self-rescuers, these systems can generate real-time alerts, thereby significantly reducing the risk of accidents and enhancing overall miner safety.

The accurate detection and identification of key elements in the heading face—including drill pipes, drill rigs, miner-drill pipe interactions, miners, safety helmets, and compressed oxygen self-rescuers—are fundamental to anomaly detection and risk mitigation in coal mining operations. However, current monitoring methods face significant challenges. In traditional coal mining production, professional monitoring personnel rely on real-time video surveillance to oversee operations and detect anomalies[5]. This approach has notable limitations: prolonged monitoring can lead to visual fatigue, reducing the accuracy of human detection and increasing the likelihood of overlooked anomalies. Furthermore, manual monitoring lacks consistency and scalability, making it difficult to maintain high detection accuracy across multiple work shifts and complex drilling environments. Given these limitations, artificial intelligence (AI)-based monitoring systems offer a promising alternative by enabling automated, real-time anomaly detection. However, the effectiveness of AI-driven object detection models is highly dependent on the availability and quality of training datasets[6]. Existing datasets for underground mining environments are either too general or lack sufficient diversity in drill pipe work scenarios, leading to poor generalization when applied to real-world heading face conditions. Moreover, many publicly available datasets fail to capture the dynamic interactions between miners and drilling equipment, which are critical for safety monitoring and operational efficiency.

Specific datasets have promoted the innovation and optimization of artificial intelligence algorithms in various research and application fields and provided strong data support for solving industry-specific problems[7–9]. For example, in recent years, classic datasets in the field of computer vision, such as MNIST, CIFAR-10, and ImageNet[10–12], as well as modern datasets like COCO, Open Images, Cityscapes, and KITTI[13–16], have played a key role in advancing technologies like image classification, object detection, and semantic segmentation[17,18]. Additionally, datasets in remote sensing (such as DeepGlobe, SpaceNet)[19,20], medical imaging (such as LIDC-IDRI, BraTS)[21,22], human pose estimation (such as MPII, COCO Keypoints)[23,24], 3D reconstruction (such as ShapeNet, ModelNet)[25,26], autonomous driving (such as ApolloScape, Waymo Open Dataset)[27,28], and super-resolution (such as DIV2K, Set5)[29,30] have provided extensive benchmarks and resources for various computer vision algorithms, promoting cross-disciplinary research and applications. Datasets in the field of natural language processing, such as GLUE[31], SQuAD[32], OpenSubtitles[33], MultiNLI[34], and SWAG[35], have played a key role in advancing technologies like text classification, speech comprehension, named entity recognition, language modeling, dialogue systems, and reasoning[36–38].

In 2023, the emergence of the open dataset for intelligent identification and classification of longwall mining face (DsLMF+) filled the gap in underground coal mine datasets[39]. This dataset includes multiple categories: miners, large coal, towline, hydraulic support guard plate and so on. However, the heading face focuses on safety and efficiency issues during the drilling process, while the longwall mining face is more concerned with broader mining operations. Although the longwall mining face dataset provides valuable resources for coal mine safety research, the uniqueness and specialization of the heading face require the development of dedicated datasets. Developing a dataset for drilling work scenarios can not only fill the gaps in existing work and provide specialized research resources for the safety and efficiency of drilling operations but promote the development of solutions to specific risks and challenges in coal mine drilling operations. To address these challenges, this study introduces the DsDPM 66 dataset, specifically designed for intelligent monitoring and anomaly detection in drilling work scenarios. Figure 1 illustrates the comprehensive process of this study along with the construction of the dataset. This dataset comprises 105,096 meticulously labeled images, extracted from surveillance footage covering 66 different monitoring perspectives in underground drilling operations. The dataset's diversity and precision annotations aim to bridge the gap in existing research by providing high-quality, scenario-specific data for training AI models. The primary detection categories include drill pipes, drill rigs, miner-drill pipe interactions, miners, safety helmets, and compressed oxygen self-rescuers. By leveraging DsDPM 66, AI-driven monitoring systems can significantly improve real-time risk detection, enhance operational safety, and optimize coal mining productivity.

Based on the analysis of common object detection dataset formats and production methods, this dataset was created by personnel familiar with coal mining operations. The dataset was annotated in YOLO format using Labelimg software, making it convenient for use in YOLO series object detection networks[40]. The dataset's applicability was enhanced by utilizing a label format conversion script for converting annotations to the COCO

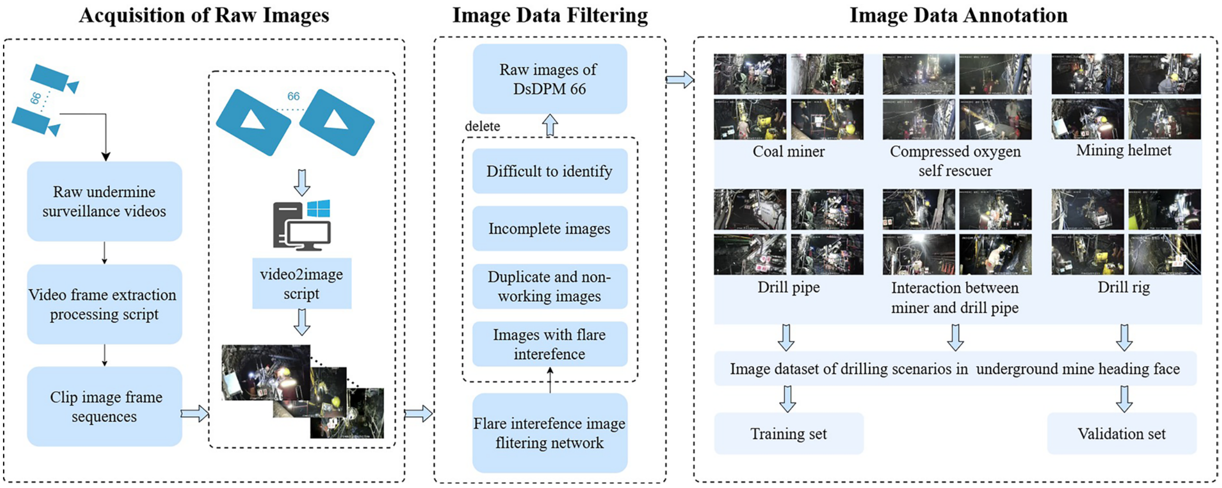


Fig. 2 General overview of the DsDPM 66 dataset construction process.

format. This modification enabled the dataset to be applied in COCO object detection methodologies. In addition to COCO and YOLO formats, other label formats can be achieved through conversion scripts.

The DsDPM 66 dataset consists of 105,096 images, meticulously captured from surveillance videos of multiple drilling operation scenes across 66 distinct sites. This dataset comprehensively documents the various dynamics and details of the drilling process. By providing both training and validation data for the development of intelligent monitoring systems, it holds substantial research value in applications related to intelligent monitoring and target detection in coal mining operations. Furthermore, the dataset aids in optimizing work processes, enhancing drilling efficiency, and improving the effectiveness of miners' personal protective measures, thereby significantly contributing to personnel safety in operational environments and facilitating the intelligent advancement of coal mining practices.

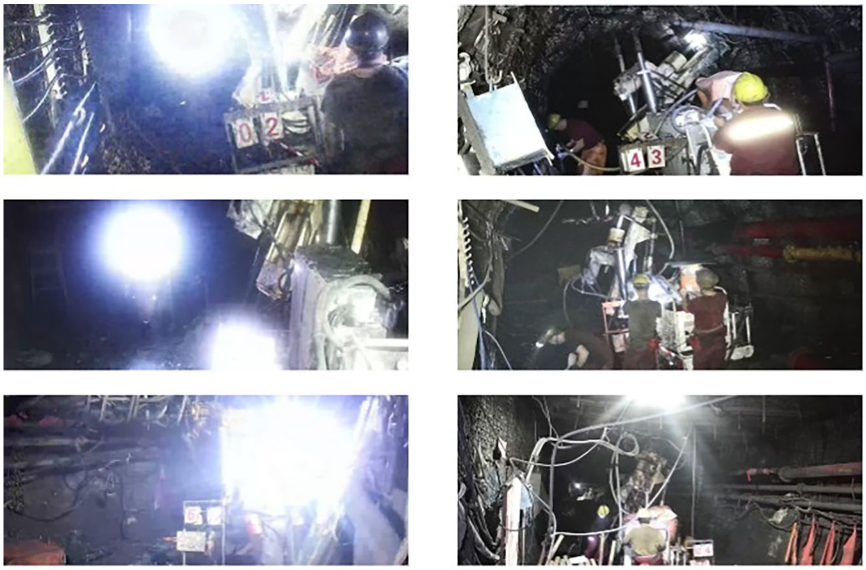
Method

The construction process of the underground coal mining heading face image dataset DsDPM 66 is shown in Fig. 2. It mainly consists of the following three steps[41]: (1) Acquisition of raw images; (2) Image filtering; (3) Image data annotation.

Acquisition of raw images. The original monitoring videos of the underground coal mining heading face were provided by Pingdingshan Coal Mine and Shaanxi Shenmu Coal Industry. We first screened and classified the monitoring videos based on different target objects, excluding those with significant obstructions from camera angles or dim lighting conditions. The DsDPM 66 dataset developed in this study consists of six categories: drill pipe, drill rig, interaction between miner and drill pipe, coal miner, mining helmet, and compressed oxygen self-rescuer. Since some of the original video data did not contain the target categories to be annotated, such as drill pipe, drill rig, interaction between miner and drill pipe, coal miner, and compressed oxygen self-rescuer, these videos were also excluded. Consequently, we obtained 66 video segments from various camera angles and shooting times, which were subsequently sorted and numbered. A 'video2image' cropping script was then utilized to extract images from the 66 monitoring videos at a rate of two frames per second, generating the raw image source for the DsDPM 66 dataset.

Image data filtering. This step involves the selection of raw images for the DsDPM 66 dataset. The DsDPM 66 dataset compiled in this study primarily includes six categories: drill pipe, drill rig, interaction between miner and drill pipe, coal miner, mining helmet, and compressed oxygen self-rescuers. Some raw images in the dataset may lack the required target categories, contain incomplete target information, or exhibit poor image quality. Such deficiencies can hinder the training and inference of AI algorithms, complicating target recognition. During the image dataset production process, we manually and automatically eliminate these potentially anomalous images.

The underground coal mining heading face is generally located in coal seams several hundred to thousands of meters below the surface, so there is no natural light. The lighting in the drilling work environment is dim, and the headlamps worn by the workers move with them, causing the light source to be unstable, resulting in fluctuations in the brightness of the monitoring video. Especially when the headlamps on the miners' safety helmets are pointed directly at the camera, it causes flare interference in the images captured by the camera. In contrast, usable clear images are shown in Fig. 3. Images exhibiting flare interference must be removed from the dataset. To address this, we developed an automatic flare interference filtering model based on PyTorch[42] and ResNet[43] to filter out low-quality images with flare interference automatically. In this project, we specifically chose low-quality images affected by glare and high-quality images with normal lighting from the raw image dataset. These images were used to create a dataset for training and validating a network model designed to eliminate glare interference. We provide this automatic flare interference filtering model and its dataset.



(a) (b)

Fig. 3 Images with flare interference to be filtered out: **(a)** images with flare interference; **(b)** images without flare interference.

In addition to the above situations, images containing abnormal data that still need to be removed include the following cases: 1) When the drilling work scene is severely affected by environmental factors such as dust and insufficient illumination, it is difficult to collect images of drill pipes, drill rigs, interactions between miner and drill pipe, coal miners, mining helmets, and compressed oxygen self rescuers. 2) Due to the limited field of view of the camera and obstruction by objects or personnel, the target collection in the images is incomplete, resulting in images that only contain partial features of the targets. 3) When the drill rig platform stops working, the camera still collects images, resulting in duplicate images in the collected raw dataset. 4) The target objects in the underground video collection are in motion. When converting these videos to images, a reasonable frame rate should be used according to different movement speeds. However, if the targets move too fast, the images obtained from the video conversion will inevitably be blurry. 5) Due to the underground environment and the distance between the target and the camera, it is challenging to distinguish distant target objects from other equipment. We manually reviewed the clear images processed by the flare interference automatic filtering model and removed those containing these abnormal data. The remaining images are used as the annotated images for the DsDPM 66 dataset.

Image data annotation. The filtered raw image dataset was annotated and labeled using LabelImg software. We provide an official open-source download link for the LabelImg software here (http://github.com/heartexlabs/labelImg). The LabelImg tool was used to annotate the training and validation sets in YOLO format. Additionally, we converted the YOLO dataset to COCO format using a script file, which we have retained. Figure 4 shows the label annotations for drill pipes, drill rigs, interactions between miner and drill pipe, coal miners, mining helmets, and compressed oxygen self rescuers. The labeled dataset can be used to develop or optimize image recognition systems that can automatically detect and identify critical equipment and personal protective equipment in real-time monitoring. For example, every miner entering a hazardous area should wear a safety helmet and self-rescuer. This automated detection helps promptly correct any lapses in safety equipment, significantly enhancing workplace safety. In the event of an accident, accurate image analysis can help quickly understand the context and specifics of the incident. For example, analyzing images of the accident scene to determine whether all safety equipment was correctly used and if any equipment abnormalities existed. This aids in post-incident analysis and the formulation of more effective countermeasures.

In addition to the single-labeled dataset, interactions between miners and drill pipes are categorized into two distinct actions: dismantling drill pipes and handling drill pipes. These action categories facilitate a detailed understanding of miners' specific skills and action sequences during operations. They are essential for developing more efficient training programs, assessing operational risks, and designing more user-friendly workflows. From a data science perspective, clearly distinguishing these behaviors into two separate categories allows for more accurate training and validation of deep learning models, thereby improving the accuracy and reliability of models in practical applications. This distinction is crucial for developing automated monitoring systems and decision-support tools aimed at improving mine safety management and operational efficiency. A more profound understanding of operations can be achieved by classifying the interactions between miners and drill pipes, which promotes advancements in technology and safety practices.

To ensure the generality and compatibility of this dataset, we collected images of drill pipes, drill rigs, interaction between miner and drill pipe, coal miners, mining helmets, and compressed oxygen self rescuers from

Fig. 4 Annotated labels for the DsDPM 66 dataset: **(a)** Coal miner; **(b)** Compressed oxygen self rescuer; **(c)** Drill pipe; **(d)** Drill rig; **(e)** Interaction between miner and drill pipe; **(f)** Mining helmet.

video clips obtained from 66 monitoring perspectives. Tables 1, 2, 3, 4, 5, 6 provide an overview of the validation and training sets for coal miners, compressed oxygen self rescuers, drill pipes, drill rigs, interactions between miner with drill pipe, and mining helmets, respectively. In the DsDPM 66 dataset constructed in this work, there are 15,606 images of miners, with 12,461 images in the training set and 3,145 images in the validation set; 19,786 images of self-rescuers, with 15,809 images in the training set and 3,977 images in the validation set; 26,842 images of drill pipes, with 21,448 images in the training set and 5,394 images in the validation set; 14,119 images of interactions with drill pipes, with 11,272 images in the training set and 2,847 images in the validation set; 15,311 images of drill rigs, with 12,224 images in the training set and 3,087 images in the validation set; and 13,432 images of mining safety helmets, with 10,719 images in the training set and 2,713 images in the validation set. The training and validation sets for each category in the DsDPM 66 dataset are divided in a 4:1 ratio. Choosing a 4:1 ratio[44] for the training and validation sets helps ensure the model learns sufficient data while allowing for effective performance validation through an independent dataset.

Data Records

The DsDPM 66 dataset is publicly available and published on the Figshare data repository[45]. Figure 5 shows the structure of the dataset. It comprises six folders, each named after a category in the DsDPM 66 dataset. These folders are named 'coal_miner', 'compressed_oxygen_self_rescuer', 'mining_helmet', 'drill_pipe', 'drill_rig' and 'interaction_between_miner_and_drill_pipe'. Each folder contains the original images, YOLO format label files,

Collection statistics	Characteristics	Training set						Validation set					
	Data size(MB)	4136.96						1044.48					
	Data distribution	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count
		001	109	023	239	045	150	001	28	023	60	045	38
		002	194	024	124	046	169	002	49	024	32	046	43
		003	125	025	969	047	192	003	32	025	243	047	48
		004	180	026	194	048	249	004	45	026	49	048	63
		005	228	027	224	049	316	005	58	027	57	049	79
		006	4	028	182	050	253	006	1	028	46	050	64
		007	323	029	168	051	86	007	81	029	43	051	22
		008	579	030	299	052	164	008	145	030	75	052	41
		009	148	031	233	053	107	009	38	031	59	053	27
		010	37	032	158	054	499	010	10	032	40	054	125
		011	148	033	112	055	200	011	37	033	29	055	51
Category	Coal miner	012	242	034	135	056	57	012	61	034	34	056	15
		013	26	035	156	057	40	013	7	035	39	057	10
		014	162	036	236	058	27	014	41	036	59	058	7
		015	127	037	343	059	35	015	32	037	86	059	9
		016	28	038	210	060	138	016	7	038	53	060	35
		017	131	039	242	061	192	017	33	039	61	061	48
		018	131	040	116	062	24	018	33	040	29	062	7
		019	58	041	60	063	460	019	15	041	16	063	115
		020	154	042	172	064	159	020	39	042	44	064	40
		021	88	043	109	065	205	021	22	043	28	065	52
		022	467	044	201	066	168	022	117	044	51	066	42
		Total	12461					Total	3145				

Table 1. The summary of the training set and validation set for coal miner.

Collection statistics	Characteristics	Training set						Validation set					
	Data size(MB)	5212.16						1310.72					
	Data distribution	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count
		001	60	023	304	045	195	001	16	023	76	045	49
		002	3	024	184	046	274	002	1	024	46	046	69
		003	0	025	1668	047	100	003	0	025	418	047	25
		004	17	026	271	048	240	004	5	026	68	048	60
		005	0	027	228	049	444	005	0	027	58	049	112
		006	4	028	329	050	338	006	1	028	83	050	85
		007	9	029	145	051	0	007	3	029	37	051	0
		008	445	030	387	052	0	008	112	030	97	052	0
		009	252	031	320	053	240	009	63	031	80	053	60
		010	96	032	248	054	544	010	25	032	62	054	137
		011	168	033	256	055	1379	011	43	033	65	055	345
Category	Compressed oxygen self rescuer	012	275	034	0	056	160	012	69	034	0	056	40
		013	10	035	264	057	242	013	3	035	67	057	61
		014	16	036	325	058	80	014	42	036	82	058	21
		015	170	037	657	059	0	015	43	037	165	059	0
		016	0	038	267	060	35	016	0	038	67	060	9
		017	14	039	315	061	240	017	36	039	79	061	60
		018	184	040	171	062	0	018	46	040	43	062	0
		019	101	041	18	063	639	019	26	041	5	063	160
		020	264	042	260	064	149	020	67	042	65	064	38
		021	0	043	175	065	244	021	0	043	44	065	61
		022	728	044	282	066	96	022	182	044	71	066	24
		Total	15809					Total	3977				

Table 2. The summary of the training set and validation set for compressed oxygen self rescuer.

Collection statistics	Characteristics	Training set						Validation set					
	Data size(MB)	7311.36						1832.96					
	Data distribution	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count
Category	Dril pipe	001	22	023	747	045	288	001	6	023	187	045	73
		002	265	024	540	046	252	002	67	024	136	046	64
		003	263	025	924	047	312	003	66	025	232	047	79
		004	228	026	468	048	104	004	57	026	118	048	26
		005	129	027	632	049	386	005	33	027	159	049	97
		006	19	028	400	050	183	006	5	028	100	050	46
		007	1259		029	437	051	399	007	315	029	110	051 100
		008	453	030	367	052	247	008	114	030	92	052	62
		009	139	031	204	053	144	009	35	031	51	053	36
		010	120	032	323	054	300	010	31	032	81	054	75
		011	840	033	275	055	288	011	210	033	69	055	72
		012	642	034	147	056	421	012	161	034	37	056	106
		013	32	035	384	057	487	013	8	035	96	057	122
		014	439	036	412	058	147	014	110	036	104	058	37
		015	318	037	588	059	47	015	80	037	147	059	12
		016	29	038	209	060	483	016	8	038	53	060	121
		017	246	039	235	061	304	017	62	039	59	061	77
		018	534	040	392	062	153	018	134	040	98	062	39
		019	182	041	30	063	1	019	46	041	8	063	1
		020	84	042	316	064	268	020	22	042	80	064	68
		021	12	043	348	065	222	021	4	043	87	065	56
		022	810	044	196	066	373	022	203	044	50	066	94
		Total	21448					Total	5394				

Table 3. The summary of the training set and validation set for drill pipe.

and COCO format annotation files for that category. These are named images, 'YOLO_labels', and 'COCO_annotations', respectively.

The images folder contains two subfolders named 'train' and 'val'. The 'train' and 'val' folders include all original image files for their respective categories, representing the training and validation sets, respectively, with a ratio of 4:1 in the number of image files. Each image file is named with an underscore, where the first three digits represent the corresponding scene index, and the last five digits represent the sequence number within that scene. The 'YOLO_labels' folder contains two subfolders named 'train' and 'val'. These folders house the YOLO format label files corresponding to the training and validation sets of images for each category. The label files contain information such as category ID, bounding box coordinates, and bounding box dimensions. YOLO format label files are in txt format, and the file names are the same as the corresponding image file names. The 'COCO_annotations' folder contains two JSON format files named after the corresponding category. One file is the COCO format annotation file for the training set, and the other is for the validation set. These COCO format annotation files are converted from YOLO format label files using the 'txt2coco' script.

Technical Validation

The DsDPM 66 dataset consists of multiple label categories, including coal miner, mining helmet, compressed oxygen self rescuer, drill pipe, drill rig, and interaction between miner and drill pipe. Each category's image data has been meticulously selected and annotated to ensure the labels accurately reflect the mining operations. We organized a review team of three professionals with extensive coal mine fieldwork experience who were responsible for checking each image and its corresponding label in the dataset. Their main tasks are to identify and correct any possible labeling errors or omissions, including but not limited to the accuracy of the labels, the clarity of the images, and the consistency of the data. The review team conducted detailed discussions for disputed labels, such as specific categories of interaction between the miner and drill pipe and the operational status of drill rigs. When necessary, final labels were determined through collective voting, ensuring that each decision was based on the collective expertise and experience of the team. To further ensure the dataset's applicability and the rigor of scientific research, we conduct regular data quality assessments. This includes rechecking label consistency and updating and improving the annotation guidelines. We will provide a complete dataset description document, including the technical parameters of image acquisition, the detailed annotation process, explanations of the team decision mechanisms, and any possible dataset limitations. This ensures research transparency and allows other researchers to replicate or extend our work. Through these measures, the DsDPM 66 dataset aims to provide a highly reliable and scientifically effective data resource for studying miner behavior and safety technologies.

	Characteristics	Training set						Validation set					
	Data size(MB)	3645.44						920.00					
Collection statistics	Data distribution	Monitor view	Labe0 Count	Label1 Count	Monitor view	Labe0 Count	Label1 Count	Monitor view	Labe0 Count	Label1 Count	Monitor view	Labe0 Count	Label1 Count
Category	Interaction between miner and drill pipe	001	0	0	034	68	63	001	0	0	034	11	22
		002	8	124	035	71	50	002	2	31	035	19	12
		003	0	0	036	221	189	003	0	1	036	47	56
		004	6	102	037	70	273	004	0	27	037	26	60
		005	21	129	038	69	93	005	4	34	038	18	23
		006	0	0	039	172	115	006	0	0	039	42	30
		007	31	153	040	93	84	007	16	31	040	24	21
		008	0	0	041	6	0	008	0	0	041	2	0
		009	0	0	042	94	71	009	0	0	042	26	16
		010	0	0	043	44	81	010	1	0	043	12	20
		011	0	2	044	65	133	011	0	1	044	21	29
		012	0	3	045	87	112	012	0	1	045	18	32
		013	0	0	046	90	98	013	0	0	046	24	24
		014	0	0	047	81	146	014	0	1	047	18	39
		015	78	33	048	109	123	015	17	11	048	27	32
		016	37	11	049	130	126	016	10	2	049	34	31
		017	95	45	050	143	163	017	28	7	050	38	39
		018	91	58	051	6	219	018	23	15	051	0	57
		019	35	29	052	106	158	019	8	9	052	32	34
		020	209	48	053	107	69	020	53	12	053	23	21
		021	73	11	054	188	167	021	16	5	054	44	45
		022	318	142	055	155	173	022	81	35	055	44	38
		023	174	122	056	151	108	023	45	29	056	40	25
		024	199	141	057	72	161	024	51	35	057	15	44
		025	146	86	058	109	19	025	32	26	058	24	8
		026	169	77	059	35	13	026	42	20	059	7	5
		027	137	57	060	129	77	027	32	17	060	38	14
		028	41	143	061	89	59	028	15	31	061	17	20
		029	93	92	062	0	0	029	19	28	062	0	0
		030	20	156	063	97	101	030	9	35	063	30	20
		031	109	81	064	124	105	031	25	23	064	32	26
		032	96	142	065	247	185	032	26	34	065	57	52
		033	51	97	066	42	147	033	17	21	066	11	37
		Total	11272					Total	2847				

Table 4. The summary of the training set and validation set for interaction between miner and drill pipe. Among which, Label0 represents handling drill pipe, Label1 represents dismantling drill pipe.

The DsDPM 66 dataset offers YOLO and COCO dataset formats, facilitating its application in top-ranked object detection neural networks currently in use. In order to assess the feasibility of the curated dataset, this study selected three top deep learning networks from the COCO object detection ranking list, namely YOLOv8[46], Swin-Transformer[47] and DETR[48], to conduct model training and validation. The access links for validating the dataset with YOLOv8, Swin-Transformer, and DETR are https://github.com/ultralytics/ultralytics, https://github.com/microsoft/Swin-Transformer, and https://github.com/facebookresearch/detr, respectively. The DsDPM 66 dataset was trained on a machine with Intel(R) Core(TM) i7-13700KF CPU, RTX 3090 GPU, and Windows 10 operating system. The three object detection algorithms' hyperparameters were based on the recommended default values. Some hyperparameters were modified to adapt to the dataset, including width, height, batch size, initial learning rate, and epoch. These changes were implemented based on the recommendations from the initial studies of YOLOv8, Swin-Transformer, and DETR.

For dataset validation, training, and evaluation, the categories include coal miner, compressed oxygen self escuer, drill pipe, drill rig, interaction between miner and drill pipe, and mining helmet. Table 7 shows the benchmark results of YOLOv8, Swin-Transformer, and DETR on the DsDPM 66 dataset. Figure 6 shows the performance graphs of the three models during the validation process, with mAP value curves for each object detection network model. The mAP values of the YOLOv8 detection model reached 0.943, 0.874, 0.788, 0.827, 0.971, and 0.919, respectively. The mAP values of the Swin-Transformer detection model reached 0.903, 0.799, 0.711, 0.95, 0.746, and 0.876, respectively. The mAP values of the DETR detection model reached 0.928, 0.858, 0.795, 0.964, 0.859, and 0.907, respectively. Although some of the above mAP values indicate good model performance, there is still much room for improvement in the mAP values for the categories of drill pipes,

Collection statistics	Characteristics	Training set						Validation set					
	Data size(MB)	4075.52						1034.24					
	Data distribution	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count
Category	Drill rig	001	92	023	191	045	166	001	23	023	48	045	42
		002	308	024	129	046	196	002	78	024	33	046	49
		003	231	025	923	047	126	003	58	025	231	047	32
		004	275	026	162	048	149	004	69	026	41	048	38
		005	261	027	145	049	268	005	66	027	37	049	68
		006	2	028	157	050	104	006	1	028	40	050	27
		007	547	029	86	051	125	007	137	029	22	051	32
		008	807	030	502	052	208	008	202	030	126	052	52
		009	140	031	153	053	63	009	35	031	39	053	16
		010	31	032	163	054	476	010	8	032	41	054	119
		011	81	033	152	055	252	011	21	033	38	055	64
		012	136	034	238	056	70	012	35	034	60	056	18
		013	62	035	129	057	105	013	16	035	33	057	27
		014	92	036	152	058	44	014	24	036	38	058	11
		015	164	037	359	059	36	015	42	037	90	059	10
		016	54	038	164	060	18	016	14	038	42	060	5
		017	143	039	177	061	124	017	36	039	45	061	31
		018	132	040	84	062	170	018	33	040	21	062	42
		019	50	041	48	063	421	019	13	041	13	063	106
		020	140	042	160	064	25	020	36	042	40	064	7
		021	140	043	100	065	169	021	35	043	25	065	43
		022	343	044	212	066	92	022	86	044	53	066	24
		Total	12224					Total	3087				

Table 5. The summary of the training set and validation set for drill rig.

Collection statistics	Characteristics	Training set						Validation set					
	Data size(MB)	3491.84						883.00					
	Data distribution	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count	Monitor view	Count
Category	Ming helmet	001	84	023	182	045	149	001	21	023	46	045	38
		002	297	024	131	046	194	002	75	024	33	046	49
		003	201	025	916	047	125	003	51	025	229	047	32
		004	162	026	160	048	139	004	41	026	40	048	35
		005	266	027	118	049	226	005	67	027	30	049	57
		006	2	028	153	050	191	006	1	028	39	050	48
		007	215	029	55	051	104	007	54	029	14	051	27
		008	624	030	491	052	180	008	157	030	123	052	45
		009	87	031	161	053	92	009	22	031	41	053	24
		010	8	032	128	054	388	010	3	032	33	054	98
		011	81	033	144	055	510	011	21	033	36	055	128
		012	124	034	222	056	72	012	32	034	56	056	18
		013	57	035	104	057	107	013	15	035	26	057	27
		014	92	036	127	058	44	014	23	036	32	058	11
		015	120	037	325	059	30	015	31	037	82	059	8
		016	34	038	152	060	17	016	9	038	38	060	5
		017	86	039	173	061	119	017	22	039	44	061	30
		018	129	040	83	062	120	018	33	040	21	062	30
		019	27	041	12	063	206	019	7	041	4	063	52
		020	122	042	64	064	56	020	31	042	17	064	15
		021	68	043	111	065	173	021	17	043	28	065	44
		022	308	044	182	066	89	022	78	044	46	066	23
		Total	10719					Total	2713				

Table 6. The summary of the training set and validation set for mining helmet.

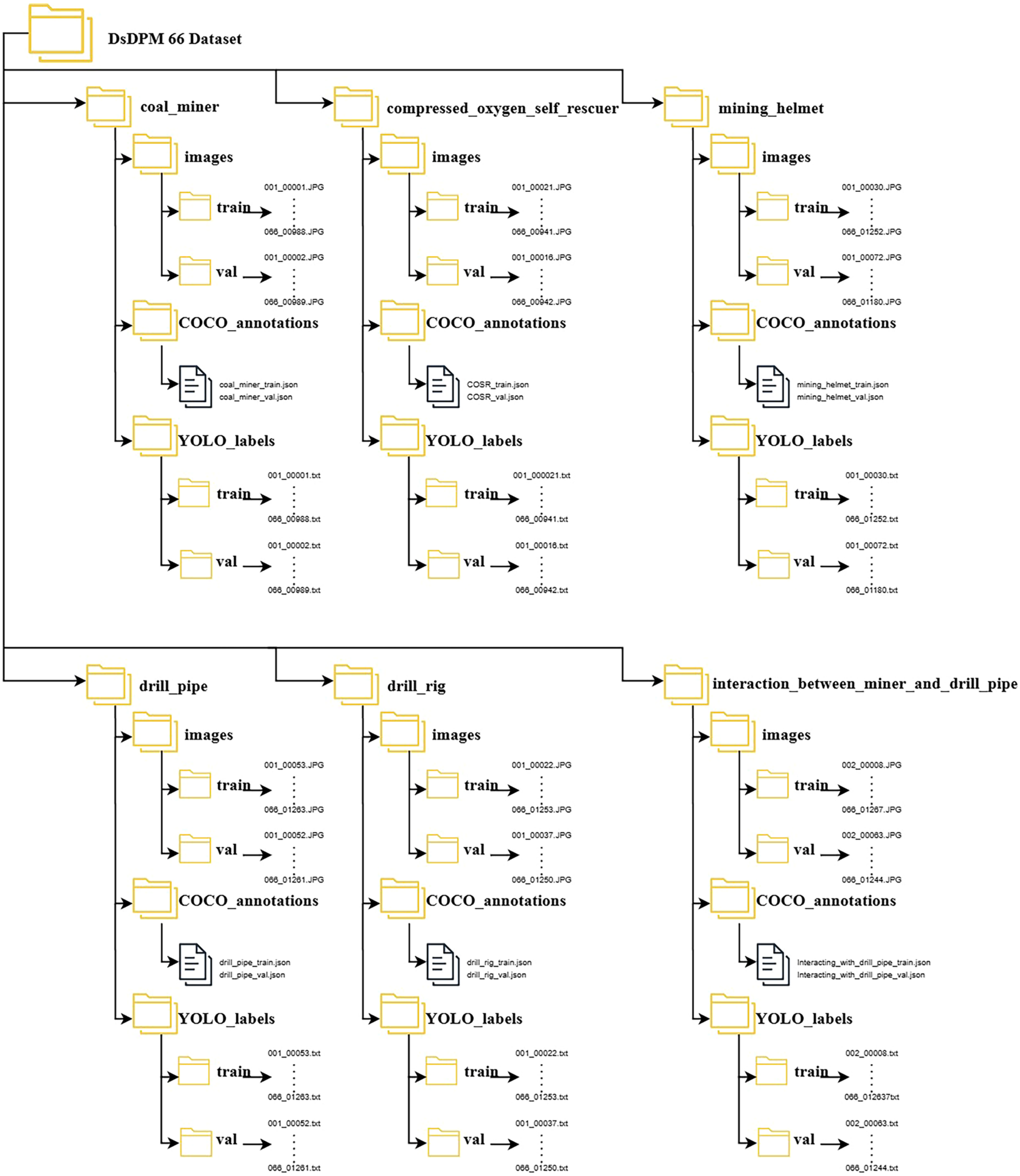


Fig. 5 DsDPM 66 dataset file folder structure, files and naming.

Category of DsDPM	YOLOv8				Swin-Transformer			DETR			
	P	R	AP0.5	AP0.5:0.95	AP0.5	AP0.75	AP0.5:0.95	AP0.5	AP0.75	AP0.5:0.95	AR0.5:0.95
Coal miner	0.984	0.992	0.993	0.943	0.989	0.967	0.903	0.988	0.964	0.901	0.928
Compressed oxygen self rescuer	0.984	0.989	0.994	0.874	0.959	0.924	0.799	0.987	0.939	0.812	0.858
Drill pipe	0.983	0.982	0.991	0.788	0.977	0.823	0.711	0.983	0.807	0.711	0.795
Drill rig	0.999	0.999	0.995	0.971	0.990	0.990	0.95	0.990	0.987	0.938	0.964
Interaction between miner and drill pipe	0.947	0.949	0.973	0.827	0.945	0.861	0.746	0.922	0.816	0.719	0.859
Mining helmet	0.997	0.989	0.994	0.919	0.989	0.968	0.876	0.988	0.977	0.881	0.907

Table 7. Benchmark of YOLOv8, Swin-Transformer and DETR performed on coal miner, compressed oxygen self rescuer, drill pipe, drill rig, interaction between miner and drill pipe and mining helmet in the DsDPM 66 dataset.

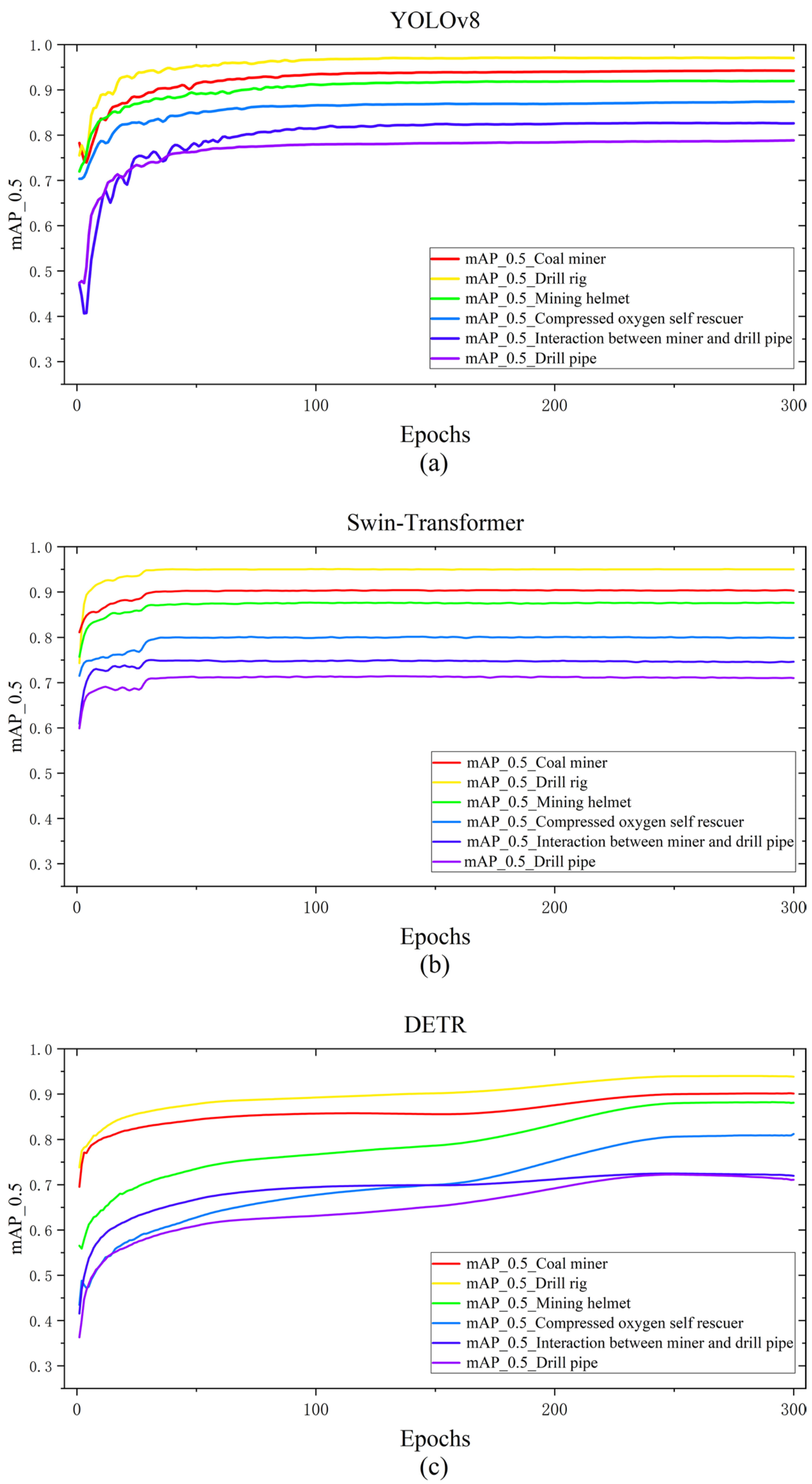


Fig. 6 Validation curves of mAP0.5 for the DsDPM 66 dataset for each category on models YOLOv8, Swin-Transformer and DETR.

compressed oxygen self-rescuers, and interaction between miner and drill pipe. This is due to the complex, low-light environment of the underground drill pipes and the inherent difficulty in recognizing these objects. Therefore, the DsDPM 66 dataset helps more researchers overcome these problems using this dataset. Figure 7 shows the results of object detection and pose estimation tasks on the models for randomly selected images from six target categories in the DsDPM 66 dataset: coal miners, compressed oxygen self rescuers, drill pipes, drill rigs, interaction between miner and drill pipe, and mining helmets. The detection effectiveness and accuracy demonstrate the reliability and practicality of the DsDPM 66 dataset.

Fig. 7 The validation of the DsDPM 66 datasets by using of the models YOLOv8, Swin-Transformer and DETR. **(a)** Coal miner; **(b)** Compressed oxygen self rescuer (COSR); **(c)** Drill pipe; **(d)** Drill rig; **(e)** Interaction between miner and drill pipe and human pose estimation map; **(f)** Mining helmet.

One of the primary challenges in underground drilling environments is the prevalence of poor lighting conditions, which significantly impacts the accuracy of object detection models. To enhance algorithm performance in low-light conditions, several strategies can be employed, including image preprocessing methodologies such as histogram equalization and Retinex-based methods, as well as the integration of infrared and thermal imaging to improve object detection under extreme conditions. Additionally, training deep learning models with augmented low-light data can increase their robustness to varying illumination levels. Beyond coal drilling, similar challenges, including low visibility, airborne dust, and occlusions, are prevalent in various underground mining operations, such as metal ore extraction and tunneling. By incorporating diverse monitoring perspectives and ensuring comprehensive annotations, our dataset provides a robust foundation for AI models that can

YOLOv8		Swin-Transformer		DETR	
dependency-package	version	dependency-package	version	dependency-package	version
Python	3.8.18	Python	3.8.19	Python	3.8.19
Pytorch	1.13.1	Pytorch	1.13.0	Pytorch	1.13.0
Torchvision	0.14.1	Torchvision	0.14.0	Torchvision	0.14.0
Cudatoolkit	11.6.55	Cudatoolkit	11.6.55	Cudatoolkit	11.6.55
ultralytics	8.1.34	mmcv-full	1.3.17	mmcv-full	1.3.17
opencv-python	4.9.0.80	mmdet	2.11.0	mmdet	2.11.0
scipy	1.10.1	pycocotools	2.0.7	mmpycocotools	12.0.3
pillow	10.2.0	timm	0.9.16	pycocotools	2.0.7
matplotlib	3.7.5	scipy	1.10.1	numpy	1.23.5

Table 8. Dependency packages required to train the DsDPM 66 dataset with YOLOv8, Swin-Transformer and DETR.

be adapted to different underground scenarios, thereby enhancing safety and operational efficiency across the mining industry.

While the DsDPM 66 dataset serves as a valuable resource for underground coal mine safety monitoring, it is important to acknowledge its limitations, including restricted environmental diversity, category imbalance, and challenges associated with extreme visibility conditions. To address these issues, future work will focus on expanding the dataset to include a broader range of mining environments and activities, ensuring a more balanced representation of object categories, and further refining preprocessing techniques. These enhancements will improve the dataset's applicability and robustness, thereby facilitating its broader adoption in mining safety and AI-driven monitoring systems.

Code availability

The DsDPM 66 dataset is publicly available on the Figshare data repository. Researchers can obtain the download link of the DsDPM 66 dataset by visiting the link (https://github.com/znrry/DsDPM-66). The image filtering network for filtering images with flare interference in underground drilling work scenes is available on the GitHub repository above. The annotation tool LabelImg can be accessed and downloaded from the official website link https://github.com/heartexlabs/labelImg. Specific usage can be referenced in the corresponding README file. In this work, the code used for training and validating the DsDPM 66 dataset employs the official open-source files released for YOLOv8, Swin-Transformer, and DETR. The codes of these three deep learning networks can be accessed via the following links https://github.com/ultralytics/ultralytics, https://github.com/microsoft/Swin-Transformer, and https://github.com/facebookresearch/detr. The required packages and their corresponding versions to run the three network models can be obtained from the 'requirements.txt' file in the link. Table 8 lists the dependency packages and versions required for training the deep learning network models YOLOv8, Swin-Transformer, and DETR on the DsDPM 66 dataset. The 'txt2coco' python file in the scripts folder of the image filtering network can be used to convert labels from YOLO format to COCO format, by visiting this link (https://github.com/znrry/DsDPM-66/blob/main/txt2coco.py). The 'video2image' python file in the scripts folder can be used to extract images from the original video data, by visiting this link (https://github.com/znrry/DsDPM-66/blob/main/video2image.py).

Received: 23 July 2024; Accepted: 1 May 2025;
Published online: 13 May 2025

References

1. Ghorbani, Y. *et al.* Moving towards deep underground mineral resources: Drivers, challenges and potential solutions. *Resources Policy* **80**, 103222 (2023).
2. Maus, V. *et al.* A global-scale data set of mining areas. *Scientific data* **7**, 289 (2020).
3. Corrigan, C. C. & Ikonnikova, S. A. A review of the use of ai in the mining industry: Insights and ethical considerations for multi-objective optimization. *The Extractive Industries and Society* **17**, 101440 (2024).
4. Yu, Y., Bai, J., Wang, X. & Zhang, L. Control of the surrounding rock of a goaf-side entry driving heading mining face. *Sustainability* **12**, 2623 (2020).
5. Du, Y., Zhang, H., Liang, L., Zhang, J. & Song, B. Applications of machine vision in coal mine fully mechanized tunneling faces: a review. *IEEE Access* (2023).
6. Zitar, R. A., Mohsen, A., Seghrouchni, A. E., Barbaresco, F. & Al-Dmour, N. A. Intensive review of drones detection and tracking: linear kalman filter versus nonlinear regression, an analysis case. *Archives of Computational Methods in Engineering* **30**, 2811–2830 (2023).
7. Sarker, S. *et al.* A comprehensive review on big data for industries: challenges and opportunities. *IEEE Access* **11**, 744–769 (2022).
8. Regona, M., Yigitcanlar, T., Xia, B. & Li, R. Y. M. Opportunities and adoption challenges of ai in the construction industry: A prisma review. *Journal of open innovation: technology, market, and complexity* **8**, 45 (2022).
9. Elahi, M., Afolaranmi, S. O., Martinez Lastra, J. L. & Perez Garcia, J. A. A comprehensive literature review of the applications of ai techniques through the lifecycle of industrial equipment. *Discover Artificial Intelligence* **3**, 43 (2023).
10. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **86**, 2278–2324 (1998).
11. Krizhevsky, A., Nair, V. & Hinton, G. Cifar-10 (canadian institute for advanced research). http://www.cs.toronto.edu/kriz/cifar.html**5**, 1 (2010).

12. Deng, J. *et al*. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, 248–255 (Ieee, 2009).
13. Lin, T.-Y. *et al*. Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, 740–755 (Springer, 2014).
14. Kuznetsova, A. *et al*. The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale. *International journal of computer vision* **128**, 1956–1981 (2020).
15. Cordts, M. *et al*. The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3213–3223 (2016).
16. Geiger, A., Lenz, P. & Urtasun, R. Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE conference on computer vision and pattern recognition*, 3354–3361 (IEEE, 2012).
17. Usamentiaga, R., Lema, D. G., Pedrayes, O. D. & Garcia, D. F. Automated surface defect detection in metals: a comparative review of object detection and semantic segmentation using deep learning. *IEEE Transactions on Industry Applications* **58**, 4203–4213 (2022).
18. Dhillon, A. & Verma, G. K. Convolutional neural network: a review of models, methodologies and applications to object detection. *Progress in Artificial Intelligence* **9**, 85–112 (2020).
19. Usmani, M., Napolitano, M. & Bovolo, F. Towards global scale segmentation with openstreetmap and remote sensing. *ISPRS Open Journal of Photogrammetry and Remote Sensing* **8**, 100031 (2023).
20. Wang, Y., Tong, L., Luo, S., Xiao, F. & Yang, J. A multi-scale and multi-direction feature fusion network for road detection from satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing* (2024).
21. Zhang, L. *et al*. Disentangling human error from ground truth in segmentation of medical images. *Advances in Neural Information Processing Systems* **33**, 15750–15762 (2020).
22. Zhang, L. *et al*. Learning from multiple annotators for medical image segmentation. *Pattern Recognition* **138**, 109400 (2023).
23. Xu, X. *et al*. Inter-image contrastive consistency for multi-person pose estimation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, 3063–3071 (2023).
24. Dong, X. *et al*. Yh-pose: Human pose estimation in complex coal mine scenarios. *Engineering Applications of Artificial Intelligence* **127**, 107338 (2024).
25. Shalma, H. & Selvaraj, P. A review on 3d image reconstruction on specific and generic objects. *Materials Today: Proceedings* **80**, 2400–2405 (2023).
26. Qi, Z. *et al*. Shapellm: Universal 3d object understanding for embodied interaction. Preprint at https://arxiv.org/abs/2402.17766 (2024).
27. Karangwa, J., Liu, J. & Zeng, Z. Vehicle detection for autonomous driving: A review of algorithms and datasets. *IEEE Transactions on Intelligent Transportation Systems* (2023).
28. Jain, S., Kumar, A., Kaushik, K. & Krishnamurthi, R. Autonomous driving systems and experiences: A comprehensive survey. *Autonomous and Connected Heavy Vehicle Technology* 65–80 (2022).
29. Musunuri, Y. R. & Kwon, O.-S. Deep residual dense network for single image super-resolution. *Electronics* **10**, 555 (2021).
30. Ma, H., Gong, B. & Yu, Y. Structure-aware meta-fusion for image super-resolution. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* **18**, 1–25 (2022).
31. Wang, B. *et al*. Adversarial glue: A multi-task benchmark for robustness evaluation of language models. Preprint at https://arxiv.org/abs/2111.02840 (2021).
32. Guven, Z. A. & Unalir, M. O. Natural language based analysis of squad: An analytical approach for bert. *Expert Systems with Applications* **195**, 116592 (2022).
33. Lison, P., Tiedemann, J. & Kouylekov, M. Opensubtitles2018: Statistical rescoring of sentence alignments in large, noisy parallel corpora. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (2018).
34. Williams, A., Nangia, N. & Bowman, S. R. A broad-coverage challenge corpus for sentence understanding through inference. Preprint at https://arxiv.org/abs/1704.05426 (2017).
35. Zellers, R., Bisk, Y., Schwartz, R. & Choi, Y. Swag: A large-scale adversarial dataset for grounded commonsense inference. Preprint at https://arxiv.org/abs/1808.05326 (2018).
36. Gienapp, L., Kircheis, W., Sievers, B., Stein, B. & Potthast, M. A large dataset of scientific text reuse in open-access publications. *Scientific Data* **10**, 58 (2023).
37. LeBel, A. *et al*. A natural language fmri dataset for voxelwise encoding models. *Scientific Data* **10**, 555 (2023).
38. Fu, J., Liu, P. & Zhang, Q. Rethinking generalization of neural models: A named entity recognition case study. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, 7732–7739 (2020).
39. Yang, W. *et al*. An open dataset for intelligent recognition and classification of abnormal condition in longwall mining. *Scientific Data* **10**, 416 (2023).
40. Vijayakumar, A. & Vairavasundaram, S. Yolo-based object detection models: A review and its applications. *Multimedia Tools and Applications* 1–40 (2024).
41. Algan, G. & Ulusoy, I. Image classification with deep learning in the presence of noisy labels: A survey. *Knowledge-Based Systems* **215**, 106771 (2021).
42. Paszke, A. *et al*. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems***32** (2019).
43. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778 (2016).
44. Cui, C., Duan, Y., Hu, H., Wang, L. & Liu, Q. Detection and classification of multiple power quality disturbances using stockwell transform and deep learning. *IEEE transactions on instrumentation and measurement* **71**, 1–12 (2022).
45. Zhao, P. *et al*. An open paradigm dataset for intelligent monitoring of underground drilling scenarios in coal mines. *figshare* https://doi.org/10.6084/m9.figshare.26135008.v2 (2024).
46. Reis, D., Kupec, J., Hong, J. & Daoudi, A. Real-time flying object detection with yolov8. Preprint at https://arxiv.org/abs/2305.09972v2 (2023).
47. Liu, Z. *et al*. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision*, 10012–10022 (2021).
48. Carion, N. *et al*. End-to-end object detection with transformers. In *European conference on computer vision*, 213–229 (Springer, 2020).

Acknowledgements

This work was supported by Aeronautical Science Foundation of China (NO.20200001012015), Ministry of Industry and Information Technology Key Laboratory of Space Optoe-lectronic Information Detection and Perception Open Project Fund(NO.NJ2023029) and ZBA Key Laboratory Open Fund (NO.6142002230102). The author would like to thank Pingdingshan Coal Mine and Shaanxi Shenmu Coal Industry who provided us with image data of the drilling scenarios and agreed to make the dataset publicly available.

Author contributions

Xichao Wang planned the production of the dataset and supervised the writing of the paper, Pengzhen Zhao collected the raw data images, trained dataset and wrote the paper, Shuainan Yu and Xiangqing Dong filtered the anomalous image data and labelled the data, Baojiang Li, Haiyan Wang and Guochu Chen reviewed and commented on the labels. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to X.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.