11th IEEE Special session on Privacy and Security of Big Data (PSBD 2024)

DIS-Mine: Instance Segmentation for Disaster-Awareness in Poor-Light Condition in Underground Mines

Mizanur Rahman Jewel, Mohamed Elmahallawy, Sanjay Madria and Samuel Frimpong

mrvfw@umsystem.edu, mohamed.elmahallawy@wsu.edu, madrias@mst.edu, frimpong@mst.edu

Overview

- Motivation
- Objectives
- Challenges
- Related works
- DIS-Mine Framework
- ImageMine Dataset
- Experimental Results

Motivation

- Disaster detection in dark underground mines is challenging, risking miners and first responders.
- Existing methods fail in low-light conditions.
- Need a solution for better situational awareness.

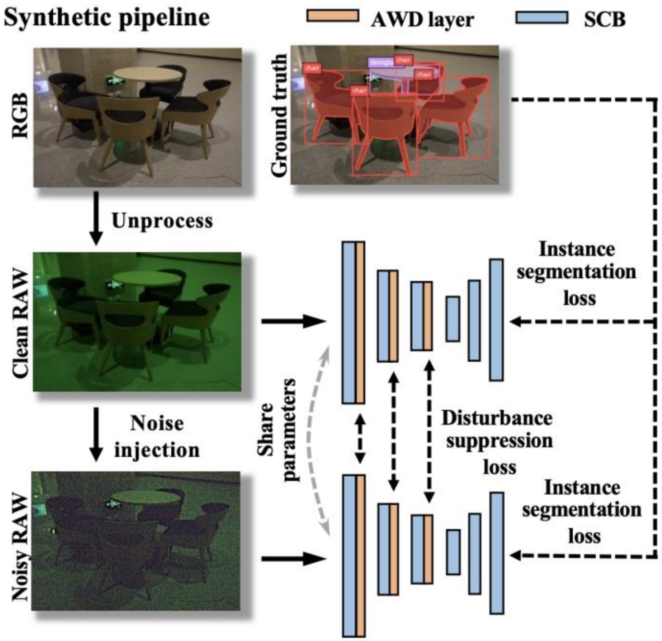
Objectives

- Develop DIS-Mine for accurate instance segmentation of disaster-affected areas in low-light underground mines.
- Create the ImageMine dataset of low-light underground mine images for model validation.
- Develop an automatic annotation pipeline for efficient labeling of mine images.

Challenges

- Low-light and Poor Visibility Conditions
- Instance Segmentation Limitations in noisy data
- Accurate Labeling and Dataset Preparation.

Related works

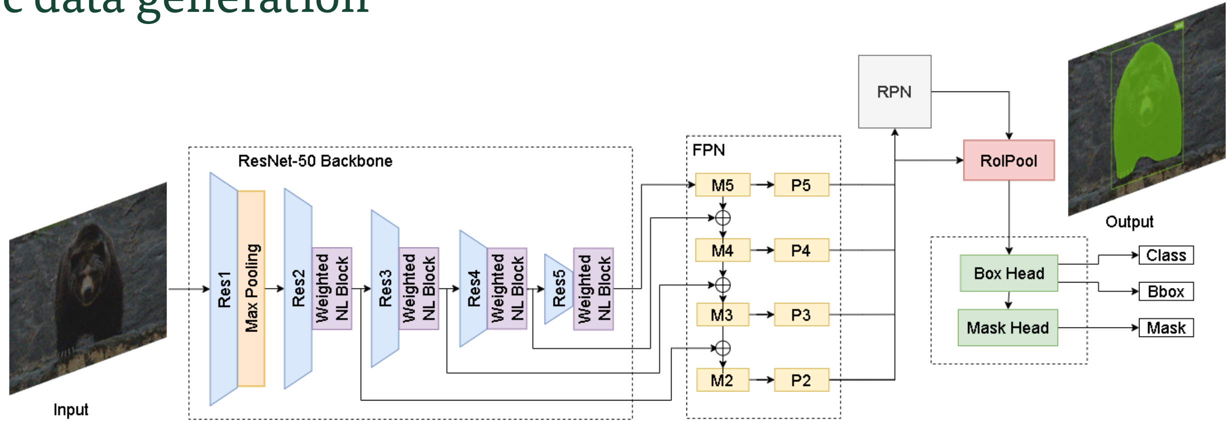


- Depending on high-bit-depth raw images
- Need clean images to train
- Synthetic data generation

Chen, L., Fu, Y., Wei, K., Zheng, D., & Heide, F. (2023). Instance segmentation in the dark. International Journal of Computer Vision, 131(8), 2198-2218.

Related works

- Primarily addresses feature-level denoising to improve segmentation performance
- color distortions and reduced contrast, are not extensively explored
- Synthetic data generation



Lin, J., Anantrasirichai, N., & Bull, D. (2024). Feature Denoising For Low-Light Instance Segmentation Using Weighted Non-Local Blocks. arXiv preprint arXiv:2402.18307.

DIS-Mine Framework

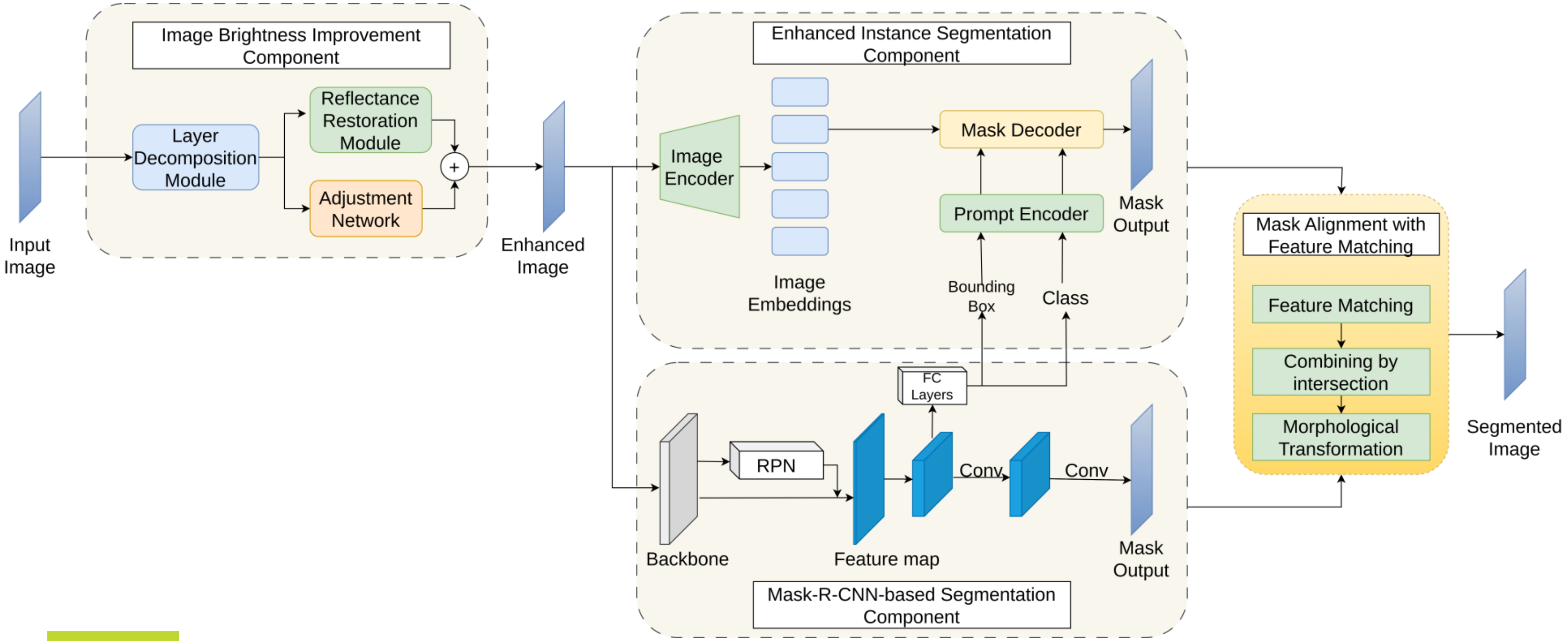


Image Brightness Improvement Component

- The image brightness improvement component integrates the KinD[1] network with DIS-Mine to enhance low-light images in the ImageMine dataset.

- The network has three modules:
  - layer decomposition (reflectance and illumination)
  - reflectance restoration
  - illumination adjustment

- Training involves calculating loss between low-light and normal image maps, restoring reflectance, adjusting illumination, and combining them to enhance brightness.

Mask R-CNN-based Segmentation Component

The mask loss is the combination of weighted dice loss and focal-loss

$$\mathcal{L}_{\text{total}} = \underbrace{\ell_{\text{class}} + \ell_{\text{box}}}_{\ell_{\text{Fast-R-CNN}}} + \underbrace{\ell_{w\text{-Dice}} + \ell_{\text{focal}}}_{\ell_{\text{enhanced-mask}}}$$

Algorithm 3: Training process of Mask R-CNN Component

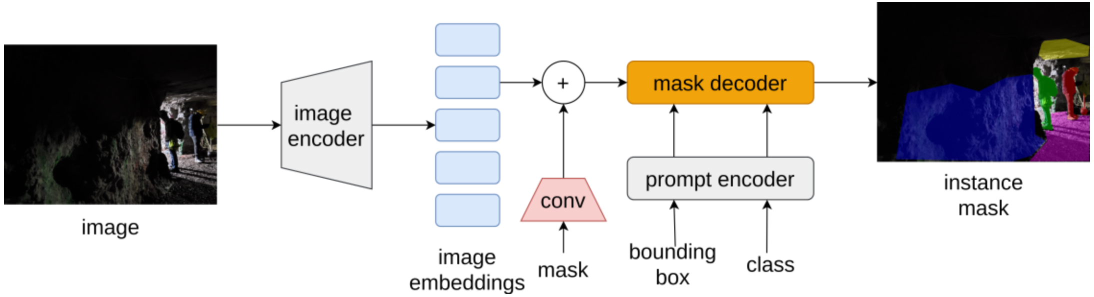
Input: Enhanced Image: I_{EN}

Output: Instance Mask: $Mask$, Class Prediction: $Class$, Bounding Box: BB

1. Initialize Mask R-CNN model
2. Add modified mask loss to the multitask loss
3. Minimize the loss and update weights
4. $Mask, Class, BB \leftarrow$ Mask R-CNN(I_{EN})
5. **Return** $Mask, Class, BB$

Instance Segmentation with SAM integration Component

Generated bounding box and class from Mask R-CNN-based segmentation component is used as prompt for SAM[4]



Mask Alignment with Feature Matching

Algorithm 4: Mask Alignment Process

Input: Enhanced Image: I_{EN}

Output: Refined-Mask: $Mask_{\text{final}}$

$Mask_2, Class, BB \leftarrow$ Mask R-CNN-based Segmentation Component(I_{EN})

$Mask_1 \leftarrow$ Instance Segmentation component$(I_i, Class, BB)$

Alignment using ORB feature-matching algorithm

$aligned\_mask_1 \leftarrow$ ORB_alignment$(Mask_1, Mask_2)$

$aligned\_mask_2 \leftarrow$ ORB_alignment$(Mask_2, Mask_1)$

$combined\_mask \leftarrow$ intersection$(aligned_mask_1, aligned_mask_2)$

$dilated\_mask \leftarrow$ dilation$(combined_mask)$

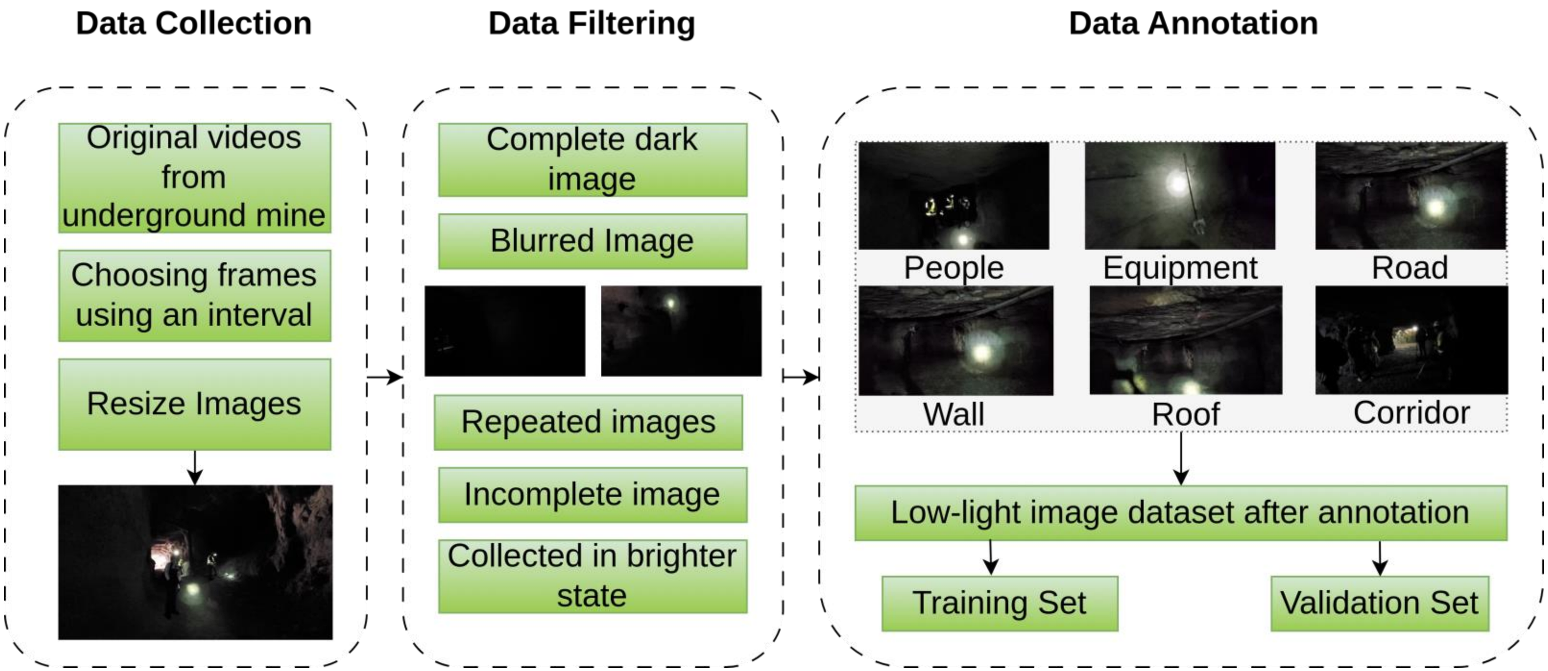
$Mask_{\text{final}} \leftarrow$ erosion$(dilated_mask)$

Return $Mask_{\text{final}}$

ImageMine Dataset

- 510 images manually annotated using VGG image Annotator
- Rest are automatically annotated

ImageMine Dataset Pipeline



Experimental Setup

- We consider a train set with six classes – people, equipment, road, wall, roof, corridor
- We evaluate the DIS-Mine framework on **ImageMine Dataset**.
- We also evaluated performance in two different dataset[2][5].
- DIS-Mine predicts the instance mask.
- We evaluate the performance using F1-score, mIoU.
- We compare with other state-of-the-art methods [2][3][4][6].

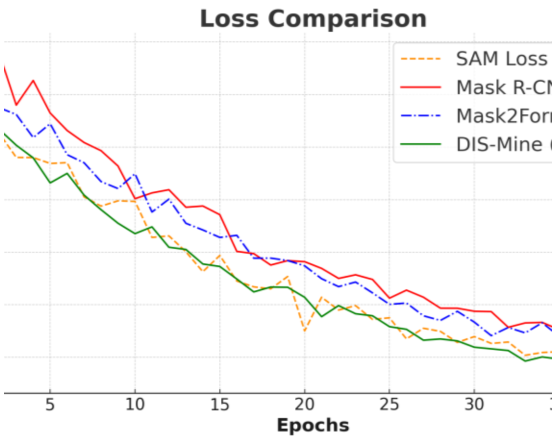
Dataset

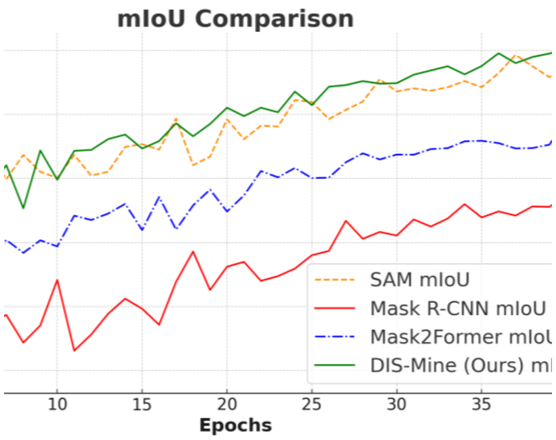
- ImageMine dataset.[our own]
- LIS-dataset[2] -
  - Pairs of low-light and normal light images
  - Total eight classes - bicycle, car, Motorcycle, bus, bottle, chair, dining table, tv.
- DsLMF+[5] -
  - Underground longwall mine images
  - Total six classes - mine personnel, hydraulic support guard plates, large coal, towlines, miners' behaviors, and mine safety helmets

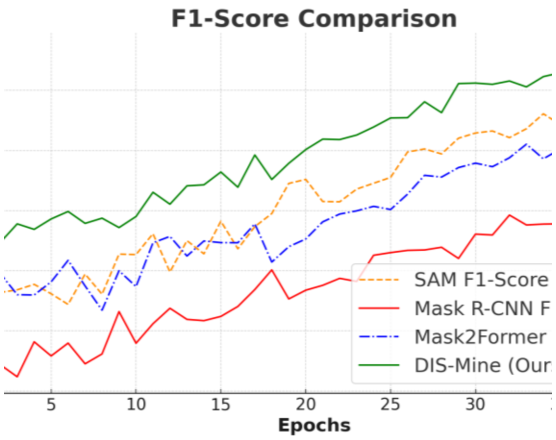
Experimental Results

Dataset	Model	F1-score	mIoU
ImageMine (Ours)	SAM	68.7%	60.0%
	Mask R-CNN	65.0%	56.0%
	Mask2Former	67.2%	58.0%
	DIS-Mine (ours)	70.2%	60.5%
LIS-dataset	SAM	61.0%	47.5%
	Mask R-CNN	58.0%	44.6%
	Mask2Former	62.0%	45.8%
	ISD	61.7%	49.8%
	DIS-Mine (ours)	63.2%	47.0%
DsLMF+	SAM	84.0%	71.0%
	Mask R-CNN	80.0%	68.0%
	Mask2Former	83.0%	72.0%
	DIS-Mine (ours)	86.0%	72.0%

Experimental Results







Experimental Results

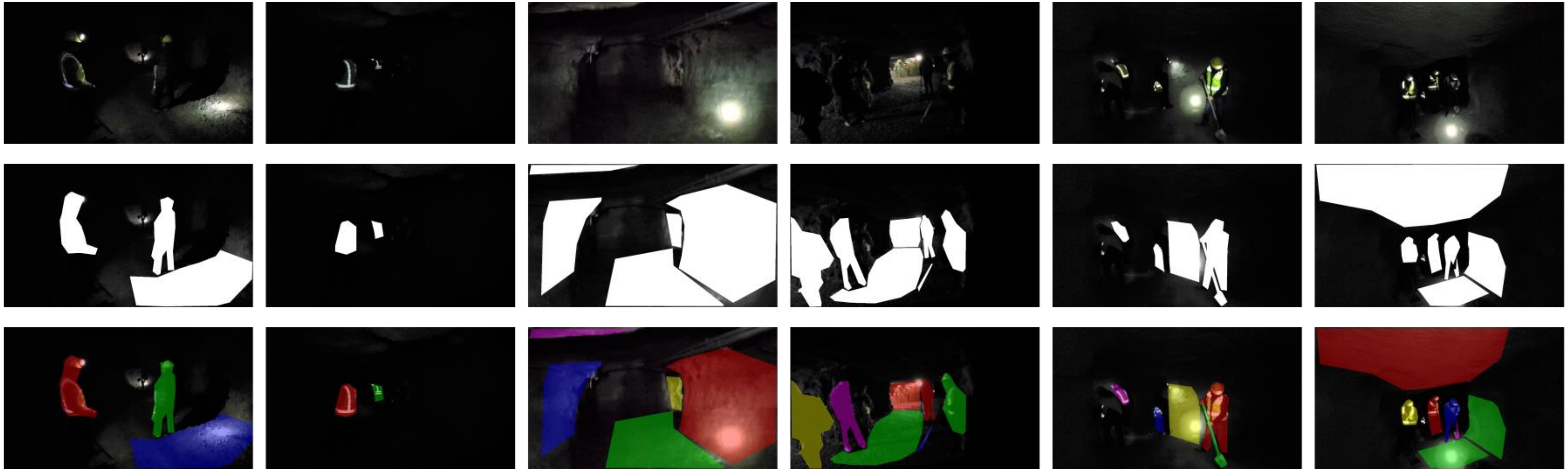


Fig: DIS-Mine prediction results on samples from our collected Image-Mine Dataset
(Input in top, ground truth in middle and generated mask in the bottom)

Experimental Results

Fig: DIS-Mine prediction results on samples from DsLMF+ Dataset
(Input in top and generated mask in the bottom)

Experimental Results

Class-Wise performance analysis for DIS-Mine on ImageMine

Surrounding consists of geometrically similar classes such as roads, walls and roofs.

Class	F1-score	IoU
People	72.6%	72.6%
Equipment	71.4%	62.1%
Corridor	88.3%	78.5%
Surrounding	64.0%	52.4%

Conclusion and Future Work

- DIS-Mine effectively segments images of underground mines in low-light conditions using advanced enhancement and segmentation techniques.
- DIS-Mine, showing superior F1-score and mIoU compared to baseline models.
- Model integration and mask alignment enhance robustness against noise and poor contrast.
- Evaluated DIS-Mine across multiple low-light datasets, proving its generalizability.
- Incorporate multimodal data (e.g., thermal, LiDAR) to improve segmentation in complex environments.

References

[1] Y. Zhang, J. Zhang, and X. Guo, “Kindling the darkness: A practical low-light image enhancer,” in Proceedings of the 27th ACM international conference on multimedia, 2019, pp. 16321640.

[2] L. Chen, Y. Fu, K. Wei, D. Zheng, and F. Heide, “Instance segmentation in the dark,” International Journal of Computer Vision, vol. 131, no. 8, pp. 2198–2218, 2023.

[3] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” in Proceedings of the IEEE international conference on computer vision, 2017, pp. 2961–2969.

[4] A. Kirillov et al., “Segment anything,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 4015–4026.

References

[5] X. Zhang, W. Yang, B. Ma, and Y. Wang, “Dslmf+: An open dataset for intelligent recognition of abnormal condition in underground longwall mining face,” 2024.

[6] B. Cheng, I. Misra, A. G. Schwing, A. Kirillov, and R. Girdhar, “Masked-attention mask transformer for universal image segmentation,” arXiv, 2021.

Acknowledgement

Funded by a grant from CDC-NIOSH


THANK YOU!