

矿井智能安全监控系统中多模态数据预警模型构建

谭小兵

乐矿能源集团(山西)矿业公司 山西 太原 030032

摘要: 随着我国煤炭工业向智能化、绿色化、安全化方向加速转型, 矿井安全生产面临更高要求。传统矿井安全监控系统多依赖单一传感器数据, 存在信息孤岛、误报率高、预警滞后等问题, 难以满足复杂地质条件下对灾害的早期精准识别需求。本文针对上述问题, 提出一种面向矿井智能安全监控系统的多模态数据预警模型。该模型融合气体浓度、微震信号、视频图像、温度湿度及人员定位等多源异构数据, 通过多模态特征提取、时空对齐与融合策略, 构建基于深度学习的联合预警架构。实验结果表明, 所提模型在瓦斯突出、顶板冒落、火灾等典型矿井灾害场景下, 预警准确率提升至99%, 误报率降低至1%以下, 平均预警提前时间达12.6分钟, 显著优于传统单模态预警方法。本研究为构建高鲁棒性、高时效性的矿井智能安全监控系统提供了理论支撑与技术路径。

关键词: 矿井安全; 多模态数据; 智能监控; 深度学习; 预警模型; 特征融合

引言

矿井作业环境复杂、风险因素多元, 瓦斯突出、顶板冒落、火灾、透水等灾害频发, 严重威胁矿工生命安全与国家能源生产稳定。据国家矿山安全监察局统计, 2023年全国煤矿共发生事故127起, 死亡213人, 其中因监测预警不及时导致的事故占比超过60%。传统矿井安全监控系统主要依赖单一物理量(如CH₄浓度、CO浓度、风速等)设定阈值进行报警, 缺乏对多因素耦合作用的综合研判能力, 导致“高误报、低预警”现象普遍存在。近年来, 物联网等技术发展使矿井感知网络完善, 形成多类型传感器的异构数据采集体系, 构成多模态数据集。但有效融合时空异步、语义异构、尺度不一的多模态信息, 构建智能模型, 是矿井安全领域核心挑战。多模态学习整合不同模态信息, 能提升模型鲁棒性与判别能力, 在多领域成效显著。不过, 矿井安全场景中, 多模态数据强噪声、低信噪比、样本不平衡, 且模态间关联复杂, 传统融合方法难以有效建模。本文聚焦矿井智能安全监控多模态预警问题, 提出基于深度神经网络的多模态数据预警模型(MDEWM)。

1 相关工作

1.1 矿井安全监控技术发展

早期矿井监控系统以模拟信号为主, 功能单一。2020年后, 随着《煤矿安全监控系统通用技术要求》(AQ6201-2019)的颁布, 数字监控系统逐步普及。近年来, 智能矿山建设推动了监控系统向“感知-分析-决策-控制”闭环演进。文献^[1]提出基于BP神经网络的瓦斯浓度预测模型, 但未考虑其他环境因素。文献^[2]利用LSTM对微震信号进行时序建模, 用于顶板稳定性评估, 但忽略

了视频与气体数据的协同作用。

1.2 多模态融合方法

多模态融合可分为早期融合(EarlyFusion)、晚期融合(LateFusion)与混合融合(HybridFusion)。早期融合在输入层拼接特征, 计算效率高但易受噪声干扰; 晚期融合在决策层集成各模态结果, 鲁棒性强但忽略模态间交互。近年来, 基于注意力机制的融合方法(如Transformer)在跨模态对齐中表现优异。比如在医疗诊断中利用Cross-Attention融合影像与文本, 提升了诊断准确率。然而, 矿井数据具有强时序性与空间关联性, 需设计专用融合架构。

1.3 深度学习在安全预警中的应用

深度学习在矿井安全预警中的应用逐渐深入。卷积神经网络(CNN)擅长处理图像与空间特征, 被广泛用于视频烟雾检测与裂缝识别; 循环神经网络(RNN)及其变体LSTM适用于时序建模, 在气体浓度预测与微震信号分析中表现良好; 而图神经网络(GNN)则能有效建模传感器间的空间拓扑关系。将矿井传感器部署视为图结构, 利用图注意力网络(GAT)进行异常检测, 但该方法未融合视频等非结构化数据, 限制了其对视觉异常(如设备冒烟、人员聚集)的感知能力。因此, 亟需一种能够统一处理结构化与非结构化多模态数据的端到端预警模型。

2 多模态数据预警模型构建

2.1 系统架构

本文提出的MDEWM模型整体采用分层架构, 包含多模态数据采集层、模态特征提取层、多模态融合层与预警决策层。在数据采集层, 系统部署气体传感器、微

震检波器、高清摄像头、红外热像仪、温湿度传感器及UWB人员定位基站,形成覆盖全矿井的异构感知网络^[3]。特征提取层针对不同数据类型设计专用编码器,以适配其内在结构特性。融合层通过时空对齐与注意力机制实现跨模态信息交互,生成统一表征。最终,预警决策层基于融合特征输出多类灾害的概率分布与综合风险评分,支持实时预警与辅助决策。

2.2 多模态数据预处理

(1) 气体与环境数据

采集频率为1Hz,经滑动窗口(窗口长60s,步长10s)生成时序片段。对CH₄、CO等关键气体进行Z-score标准化,并计算变化率、峰值等衍生特征。

(2) 微震信号

采样频率为1kHz,经带通滤波(10-200Hz)去除低频干扰与高频噪声。采用短时傅里叶变换(STFT)生成时频图,作为CNN输入。

(3) 视频图像

每秒抽取2帧,经YOLOv5进行目标检测(识别人员、设备、烟雾、裂缝等),提取ROI区域并输入ResNet-18提取视觉特征。

(4) 人员定位数据

UWB定位精度达±0.3m,记录人员坐标、移动速度、停留时间等,构建人员活动热力图。

2.3 模态特征提取网络

为充分挖掘各模态的判别性信息,本文设计了专用特征提取网络。对于气体浓度、微震能量、温湿度等时序数据,采用双向LSTM(BiLSTM)作为编码器,其前向与后向隐状态可捕捉时间序列的长期依赖关系,隐藏层维度设为128。视觉模态则采用在ImageNet上预训练的ResNet-18网络,冻结底层卷积层以保留通用特征,微调顶层以适配矿井场景,最终输出512维特征向量^[4]。对于人员定位数据,将其建模为动态图结构:节点代表人员,边权重由欧氏距离的倒数定义,输入两层图卷积网络(GCN)进行消息传递,聚合邻居信息以表征群体行为模式。

2.4 多模态融合策略

(1) 时空对齐

由于各模态采样频率不同,采用线性插值与时间戳对齐,统一为10秒时间粒度。空间上,将传感器位置与视频ROI区域进行地理映射,建立空间关联矩阵。

(2) 注意力融合机制

设计跨模态注意力模块(Cross-ModalAttention, CMA):

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

其中, Q来自视觉特征, K、V来自时序特征,实现“视觉引导时序关注”。同时引入自注意力机制增强模态内特征表达。

(3) 图神经网络融合

将各模态特征视为图节点,构建多模态异构图。节点间边由空间距离与语义相似度共同决定。通过GAT聚合邻居信息,生成全局融合特征向量 $h_{\text{fused}} \in \mathbb{R}^{256}$ 。

2.5 多任务预警决策

采用多任务学习框架,共享融合特征,分支输出不同灾害类型:

任务1: 瓦斯突出预警(一分类)

任务2: 顶板冒落预警(一分类)

任务3: 火灾预警(一分类)

任务4: 综合风险评分(回归)

损失函数为加权和:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\text{gas}} + \lambda_2 \mathcal{L}_{\text{roof}} + \lambda_3 \mathcal{L}_{\text{fire}} + \lambda_4 \mathcal{L}_{\text{risk}}$$

其中, \mathcal{L}_{cls} 采用FocalLoss缓解样本不平衡, \mathcal{L}_{reg} 采用SmoothL1Loss。

3 实验与分析

3.1 数据集与实验设置

实验数据来源于某大型煤矿2022至2024年的实际监控记录,涵盖正常运行状态12,850小时,以及23起瓦斯突出、18起顶板冒落和9起火灾事件。传感器部署包括120个气体传感器、40个微震检波器、85台高清摄像头及30个UWB基站,形成密集感知网络。数据按7:2:1划分为训练集、验证集与测试集,确保评估的客观性。模型在NVIDIA A100 GPU上训练,优化器选用AdamW,初始学习率为1e-4, batchsize设为32,训练过程采用早停策略防止过拟合。

3.2 对比方法

为验证模型有效性,选取四种代表性基线方法进行对比:

Baseline-1: 仅气体阈值报警(现行标准)

Baseline-2: LSTM(仅气体+微震)

Baseline-3: EarlyFusion(特征拼接+MLP)

Baseline-4: LateFusion(各模态独立预测后投票)

Proposed: 本文MDEWM模型

3.3 评价指标

准确率(Accuracy)

精确率(Precision)

召回率 (Recall)

F1-score

误报率 (FalseAlarmRate,FAR)

平均预警提前时间 (MeanLeadTime,MLT)

3.4 结果分析

(1) 整体性能对比 (表1)

表1 整体性能对比

方法	Accuracy(%)	F1-score	FAR(%)	MLT(min)
Baseline-1	78.2	0.65	18.7	3.1
Baseline-2	83.5	0.72	12.4	5.8
Baseline-3	86.1	0.78	9.6	7.2
Baseline-4	88.3	0.81	7.9	8.5
Proposed	99.0	0.92	<1	12.6

结果表明, MDEWM在各项指标上均显著优于基线方法, 尤其在降低误报率与提升预警提前时间方面优势突出。

(2) 消融实验

消融实验进一步证实了各模块的有效性: 移除跨模态注意力模块后F1-score下降至0.87, 移除图神经网络融合后降至0.85, 而仅使用单模态 (如仅视频) 时F1-score仅为0.68左右。在一次真实瓦斯突出事件回溯中, MDEWM在CH₄浓度仅为0.5% (远低于1%报警阈值) 时即发出预警, 提前14.3分钟, 为人员撤离争取了宝贵时间, 充分体现了模型的前瞻性与实用性。

4 讨论

4.1 模型优势与创新性

本文模型的核心优势在于实现了多源异构数据的深度融合与协同推理。不同于传统方法将各传感器孤立处理, MDEWM通过构建“眼-鼻-耳”协同感知体系, 使系统不仅能“闻”到气体异常, 还能“看”到视觉变化、“听”到岩体破裂声、“知”晓人员行为, 从而对灾害前兆形成全景式认知^[5]。动态注意力机制赋予模型可解释性, 使其决策过程可追溯; 端到端训练避免了传统流水线中特征工程与模型割裂的问题, 提升了整体性能。

4.2 局限性与现实挑战

尽管模型性能优越, 但在实际部署中仍面临若干挑战。首先, 灾害样本稀少导致数据标注成本高昂, 未来需探索半监督学习或生成对抗网络 (GAN) 进行数据

增强。其次, 模型计算复杂度较高, 需依赖边缘计算设备支持, 对矿井井下网络带宽与算力提出更高要求。此外, 不同矿区地质条件、开采工艺差异显著, 模型在新矿区的泛化能力有待验证, 需引入领域自适应技术进行迁移学习。

4.3 未来研究方向

未来工作将从三个方向展开: 一是引入大语言模型 (LLM) 解析安全规程、事故报告等文本数据, 构建“感知-认知”闭环, 使系统不仅能识别异常, 还能理解其成因并提出处置建议; 二是开发轻量化模型架构, 如知识蒸馏或神经网络剪枝, 以适配井下资源受限的边缘设备; 三是推动行业协作, 构建全国矿井安全多模态数据共享平台, 促进模型持续迭代与标准化应用。

5 结语

本文针对矿井安全监控中单模态预警的局限性, 提出了一种基于多模态数据融合的智能预警模型MDEWM。该模型通过深度神经网络有效整合气体、微震、视频、定位等多源信息, 利用时空对齐、注意力机制与图神经网络实现特征深度融合。在真实矿井数据集上的实验表明, 该模型显著提升了预警准确率与提前时间, 降低了误报率, 为构建新一代矿井智能安全监控系统提供了可行方案。未来, 随着多模态感知技术与人工智能算法的持续进步, 矿井安全生产将迈向更高水平的智能化与自主化。

参考文献

- [1]张强,李伟.基于BP神经网络的矿井瓦斯浓度预测模型[J].煤炭学报,2019,44(5):1523-1530.
- [2]Wang L, et al. Roof stability assessment using microseismic monitoring and LSTM networks[J]. International Journal of Rock Mechanics and Mining Sciences, 2021, 142: 104732.
- [3]杨晓阳.矿井智能安全生产监控系统的设计与研究[J].机械管理开发,2021,36(01):213-214+259.
- [4]陈彩华.矿井安全监控平台中的智能决策系统研究[J].贵州师范大学学报(自然科学版),2019,37(05):105-109.
- [5]李耀斌.矿井综采工作面通风安全智能监控系统应用研究[J].矿业装备,2025,(08):117-119.