

基于大模型的 化工安全风险智能识别研究

Research on Large Model-Driven Intelligent Identification for Chemical Safety Risks

杨成玮¹, 陆新建², 徐滨阳³, 周致永² (1. 友达光电(苏州)有限公司, 江苏 苏州 215021; 2. 南京科思倍信息科技有限公司, 江苏 南京 210000; 3. 中讯邮电咨询设计院有限公司, 北京 100048)

Yang Chengwei¹, Lu Xinjian², Xu Binyang³, Zhou Zhiyong² (1. AUO(Suzhou)Co., Ltd., Suzhou 215021, China; 2. Nanjing ChemCyber Technology Company Ltd., Nanjing 210000, China; 3. China Information Technology Designing & Consulting Institute Co., Ltd., Beijing 100048, China)

摘要:

在化工、能源等高危行业智能化转型中, 作业现场安全风险识别与预警成焦点, 传统检测方法难适应复杂动态场景。提出了一种基于多模态大模型的智能风险识别与预警系统, 通过结构化Prompt工程引导与企业知识库深度融合, 提升视觉大语言模型在复杂工业场景下的推理与识别能力。系统既支持视频动态目标抽取与关键帧筛选, 也可对静态场景图像进行智能分析与风险判断。该系统以自动化、智能化为核心, 显著提升了风险识别的适应性和实用性。

关键词:

化工安全; 风险识别; 智能预警; 多模态大模型; 知识库融合; Prompt工程; 人工智能

doi: 10.12045/j.issn.1007-3043.2025.06.004

文章编号: 1007-3043(2025)06-0018-05

中图分类号: TP391

文献标识码: A

开放科学(资源服务)标识码(OSID):



Abstract:

Under the intelligent transformation of high-risk industries such as chemical and energy, safety risk identification and early warning at work sites have become a focus. The traditional detection methods are difficult to adapt to complex dynamic scenarios. It proposes an intelligent risk identification and early warning system based on multi-modal large models, which enhances the reasoning and recognition capabilities of Vision-Language Models (VLMs) in complex industrial scenarios through the integration of structured Prompt engineering and enterprise knowledge bases. The system supports video dynamic object extraction and keyframe filtering, as well as intelligent analysis and risk assessment of static scene images. The system is centered around automation and intelligence, significantly enhancing the adaptability and practicality of risk identification.

Keywords:

Chemical safety; Risk identification; Intelligent early warning; Multi-modal large model; Knowledge base integration; Prompt engineering; Artificial intelligence

引用格式: 杨成玮, 陆新建, 徐滨阳, 等. 基于大模型的化工安全风险智能识别研究[J]. 邮电设计技术, 2025(6): 18-22.

0 前言

化工、能源等高危行业安全风险面临严峻挑战, 其园区化、规模化特征对智能安全监测系统提出了更高要求^[1]。传统人工巡检主观性强, 效率低下, 而单一视觉方法(如目标检测)虽提升了自动化水平^[2-3], 但在深度理解场景语义、进行复杂逻辑推理方面存在

固有局限, 难以有效应对复杂动态的作业现场风险。高危行业 AI 安全监管的迫切需求已得到政策层面的关注^[4]。行业亟需更智能、更通用的风险识别新范式。近年来, 具备强大跨模态理解与推理能力的视觉大语言模型(VLM)为此提供了创新路径。本文聚焦高危行业特殊作业场景, 提出一种基于VLM并融合企业知识库的智能风险识别与预警系统, 旨在通过深度语义理解和知识增强, 突破传统单一模态识别在复杂推理和场景适应性方面的瓶颈。该系统通过结构化Prompt

收稿日期: 2025-04-16

工程与知识增强,力求提升风险识别的广度、深度与实用性,为行业智能安全管理探索具备高自动化水平和强场景适应性的新方案。

1 多模态大模型与风险识别

1.1 多模态大模型核心能力

多模态大模型是深度学习领域的重要进展,它能够同时处理和理解文本、图像、视频等多种类型的数据。与仅处理单一类型数据的模型相比,其核心优势在于具备强大的跨模态理解能力,能够捕捉不同模态信息间的深层语义关联。工业场景下多模态数据的融合与实时风险推理框架是当前研究的热点^[5]。特别是视觉大语言模型(VLM),它通过联合训练视觉编码器和大型语言模型,能够完成复杂的视觉问答、图像描述和场景推理任务。此外,大型模型普遍展现出的零样本(Zero-shot)或少样本(Few-shot)学习能力,使其在面对标注数据稀疏或风险类型多变的工业场景时,具有更好的适应性和泛化潜力。

1.2 多模态大模型赋能风险识别

传统基于计算机视觉的风险识别方法(如目标检测、行为识别等)虽然在特定任务上取得了成功,但在应对化工安全这类复杂场景时,往往因缺乏深度语义理解和知识融合能力而受限。多模态大模型则能有效克服这些局限,其赋能风险识别的关键优势体现在如下几个方面。

a) 跨模态融合理解。多模态大模型能够整合分析来自视频监控的视觉信息、DCS/PLC系统的设备参数(时序数据)、作业票或安全规程(文本信息)等多源异构数据,实现对风险场景更全面的感知和判断,突破了单一数据源分析的局限性。

b) 深度语义理解与推理。多模态大模型不仅能识别物体或行为,更能深入理解其在特定工业语境下的含义和潜在风险,实现基于上下文和常识的复杂逻辑推理,这是传统基于模式匹配或简单规则的方法难以企及的。知识增强的视觉语言模型在工业场景适配方面展现出良好效果^[6]。

c) 灵活的适应性与扩展性。得益于零/少样本学

习能力,系统能够快速适应新的风险类型或场景变化,相较于需要针对特定任务进行大量数据标注和模型迭代的传统方法,大大降低了维护和扩展成本。针对工业危险场景的VLM基准测试也验证了其潜力^[7]。

d) 知识增强与可解释性。通过结合检索增强生成(RAG)等知识融合技术,系统可动态查询企业安全知识库(包含规程、标准、案例等),将相关领域知识融入推理过程,弥补了通用大模型在垂直领域专业知识上的不足,不仅提高了判断的专业性和准确性,还能作为风险判定提供明确的依据,增强结果的可解释性。

这种融合了感知、推理和知识的能力,使得多模态大模型特别适合应对化工行业中需要综合判断、依赖领域知识的复杂安全风险识别任务,有望推动安全管理向更智能化的方向发展。

2 化工风险智能识别系统设计

2.1 系统总体架构

基于多模态大模型的化工安全风险智能识别与预警系统总体架构如图1所示,该系统构建了从数据采集、智能分析到风险预警的完整闭环。系统包含4个核心层次:数据采集与预处理层、特征提取与场景理解层、风险识别与知识融合层以及结果生成与预警发布层。

a) 数据采集与预处理层负责接入多源异构数据,包括监控摄像头视频流、移动端采集的现场图片、DCS/PLC系统的设备参数等。针对不同数据源,该层执行格式转换、噪声过滤、时序对齐等预处理,为后续分析提供标准化的数据流。

b) 特征提取与场景理解层处理预处理后的数据,对于视频流,可选择性地采用YOLOv8等高效目标检测模型进行初步筛选和关键帧抽取;对于静态图像,则直接输入视觉大语言模型进行处理。此层的核心目标是对输入场景形成初步理解,识别出场景中的关键元素与潜在风险点。

c) 风险识别与知识融合层是系统的“智能大脑”,它利用精心设计的Prompt模板引导多模态大模型进行风险分析与判断。核心技术包括:

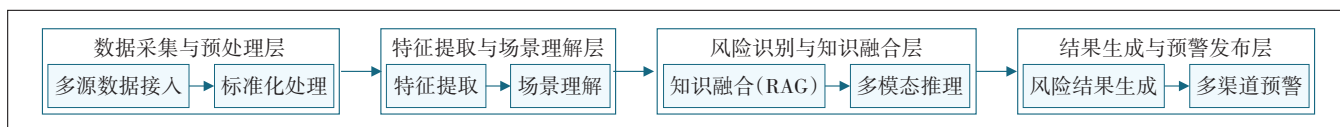


图1 系统总体架构

(a) 结构化 Prompt 工程: 根据不同风险类型构建专业化的提示模板。

(b) 检索增强生成: 实时从企业安全知识库检索相关规程、标准和案例, 将其融入 Prompt 中增强模型推理。

(c) 风险评估与分级: 根据推理结果结合预定义规则, 确定风险等级和优先级。

d) 结果生成与预警发布层将风险识别结果格式化为结构化输出(如 JSON), 输出内容包含风险描述、风险等级、具体位置、判断依据和建议措施等。同时, 系统提供多渠道预警发布(如集成企业安全管理平台、移动 APP 推送、发送短信通知等), 实现风险预警的自动化闭环。

2.2 关键技术实现

为了有效引导多模态大模型执行化工安全风险识别任务, 本系统采用了高度结构化的 Prompt 工程策略^[8]。典型的 Prompt 结构包含角色设定(如“你是一名经验丰富的化工安全工程师”)、明确的任务指令(如“分析图像中是否存在违反安全规程的操作行为”)、必要的场景描述(如“这是在 XX 区域的动火作业现场”)、通过 RAG 动态注入的相关知识(如安全规程条款)以及结构化的输出格式要求。例如, 在识别“未佩戴安全帽”风险时, Prompt 会明确角色、任务、场景背景、必须佩戴安全帽的规程依据, 并要求以 JSON 格式输出包含风险判断、描述、等级、依据和建议措施的结果。通过构建不同风险类型的 Prompt 模板库, 系统能够适应多样化的风险识别需求。

企业安全知识库与多模态大模型的融合是提升系统专业性和准确性的核心要素, 结合知识库的方法是当前 LLM 研究的重要方向^[9]。本系统采用检索增强生成(RAG)机制实现知识的有效利用。首先, 将企业安全规程、操作标准、历史事故案例、设备手册等文档进行结构化处理和向量化存储, 构建企业安全知识库。工业知识图谱与多模态大模型的协同推理是实现这一目标的关键技术^[10]。在接收到分析任务时, 系统会基于输入场景和初步理解, 从知识库中智能检索最相关的知识片段, 并将其动态注入到为该场景生成的 Prompt 中。这种方式相比模型微调, 具有知识更新便捷、专业性强、可解释性好的优势, 特别适合化工等对专业知识和安全标准高度依赖的行业。

针对化工行业多样化的数据来源和风险识别需求, 系统设计了灵活的多模态数据融合与处理流程,

支持多种输入模式。对于静态图像(如移动端拍照、固定监控截图), 可直接输入多模态大模型进行分析。对于连续的视频流, 则可先通过 YOLO 等目标检测模型进行初步筛选, 识别并抽取包含人员、特定设备或进入特定区域等关键信息的帧, 再将这些关键帧输入大模型进行深度分析, 以提高效率。此外, 系统还支持融合图像/视频信息与 DCS/PLC 等系统传输的设备实时参数(如温度、压力、气体浓度), 通过多模态大模型进行综合判断, 识别如“高温设备附近存在违规操作”等需要跨模态信息理解的复合风险。系统的具体处理流程如图 2 所示。

3 应用场景与实施路径

3.1 典型应用场景

基于多模态大模型的化工安全风险识别系统可广泛应用于化工、能源等高危行业的多种场景。

在作业行为安全管控方面, 系统聚焦于生产作业过程中的人员行为安全监管, 包括特殊作业监护、作业规范性检查等。通过监控摄像头或移动终端实时采集作业现场图像, 可识别如未佩戴防护用品、违规操作、禁区闯入等常见风险, 利用 YOLO 等模型进行危险行为检测是该系统常用的方法^[11]。例如, 在动火作业监管中, 系统能识别动火区域周围未清理可燃物、监护人员离岗使用手机等违规行为, 结合动火作业票信息, 还能判断操作是否符合许可要求, 实现了对复杂语义场景(如监护人失职)的理解。

在设备设施隐患排查方面, 系统侧重于对生产设备、安全设施的状态监测和隐患识别, 如管道泄漏、阀门异常、消防设备被遮挡等。通过定期或实时图像分析, 结合必要的设备运行参数, 实现对常见设备隐患的智能识别。例如, 在日常巡检中, 系统可识别消防通道被物料堵塞、安全阀铅封缺失或气瓶未按规定固定存放等隐患, 并能理解“通道被阻挡”这类需要结合场景和规则判断的情况。在石化等特定场景的实时异常检测已有相关研究^[12]。

在环境风险智能监测方面, 系统关注作业环境中的安全风险, 如泄漏污染、异常天气影响、外部入侵等。结合环境监测数据(如天气信息)与视觉信息, 实现对环境风险的智能预警。例如, 在危化品仓储装卸区, 系统可识别雨天作业未采取防雷措施、危化品包装破损泄漏、地面积水接触电气设备等风险, 并结合天气信息与作业活动推断操作是否合规。

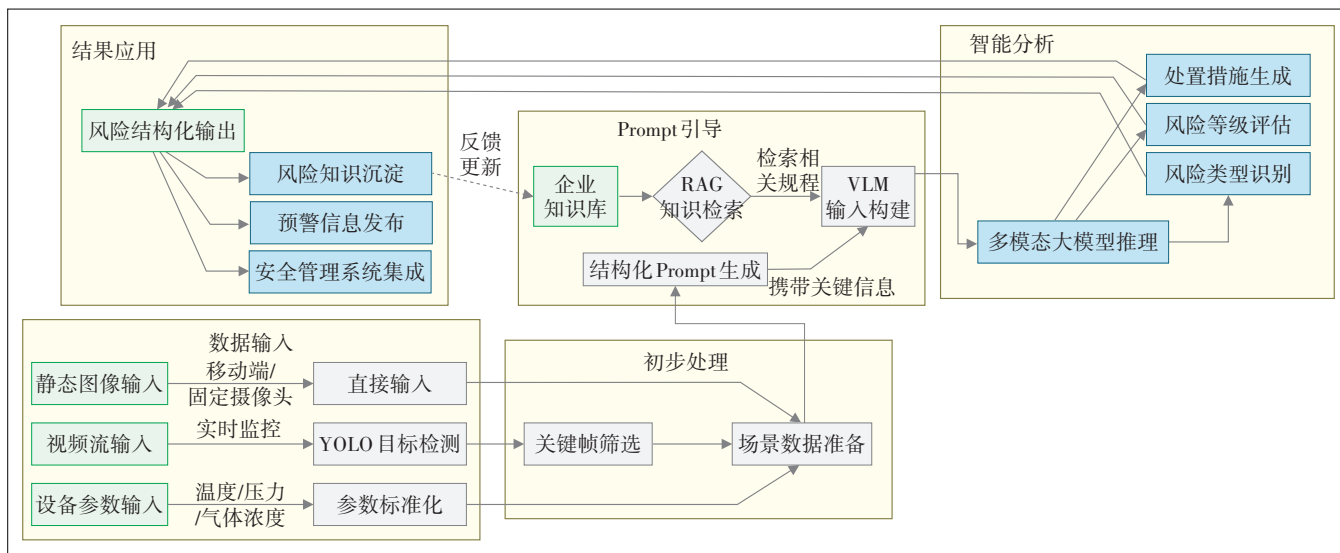


图2 多模态风险识别处理流程

在应急响应辅助决策方面,系统可在突发事件或应急演练中提供实时态势分析与处置建议。系统通过对事件现场的连续监测与分析,识别危险发展趋势(如疏散通道拥堵、救援人员防护不到位等),并结合应急预案提供处置建议,辅助指挥决策,从而提升应急响应效率。

3.2 系统实施路径

为确保系统成功落地并发挥价值,需采取聚焦应用、突出重点、迭代优化的实施策略。首先,核心能力的部署应优先聚焦于企业安全管理的痛点和高风险场景,如重大危险源监控、高风险特殊作业的实时监控以及事故多发区域的常态化巡检。根据企业基础和需求,可灵活选择部署模式:对于快速验证和中小型企业,可采用云端部署,利用云端大模型的推理能力;对于大型园区或实时性要求高的场景,则宜采用边云协同模式,在边缘侧进行初步处理(如视频分析^[13]),云端完成复杂分析。基于边缘计算的实时预警系统是重要的发展方向。同时,必须考虑与企业现有安全管理平台(如HSE、视频监控、DCS/SIS等)进行系统集成,实现数据互通和流程闭环。

其次,关键技术要素的落地是系统效能发挥的基石。这包括高质量数据准备与治理(规范接入、持续治理、保障数据时效性)、精准的企业知识库构建与维护(结构化处理、向量化存储、动态更新)、场景化的模型适配与Prompt优化(针对性优化或微调、建立模板库、持续迭代)、可靠的人机协同机制(明确AI边界、设定合理阈值、建立人工复核流程)以及严格的安全合

规保障(数据加密、脱敏、权限控制等)。

最后,系统的部署是一个持续迭代与价值评估的过程。需建立基于实际应用效果的反馈闭环机制,收集误报漏报案例,反向优化模型、Prompt和知识库。同时,建立清晰的价值评估体系,从风险管控提升(如隐患发现率提高)、管理效率优化(如响应时间缩短)、合规性与标准化提升以及经济效益(如事故预防带来的损失降低)等维度量化系统效益,不断提升应用效果和投入产出比,支撑企业安全管理水平持续提升。

4 讨论与展望

本文提出的基于多模态大模型的化工安全风险智能识别系统,通过结构化Prompt工程与企业知识库的深度融合,为解决传统风险识别方法的局限性提供了创新路径。该系统在模拟场景和部分实际场景测试中展现了其在理解复杂场景语义、推理潜在风险逻辑方面的潜力。然而,将此技术大规模应用于严苛且复杂的化工生产环境中,仍需审慎评估并克服一系列挑战,同时也要积极展望其广阔的应用前景。

4.1 关键问题与挑战

尽管前景广阔,但该系统在化工行业的深度应用仍面临若干现实挑战。

a) 成本与效率的平衡。多模态大模型的推理成本(API调用或私有化部署硬件的成本)相对较高,对于需要处理海量视频流或高频次分析的场景,需要权衡投入产出比。优化模型结构、采用边云协同架构、优先部署于高风险关键环节是降低成本的可行策略。

b) 复杂环境下的识别鲁棒性。化工现场光照变化、粉尘、水汽、遮挡以及防护装备(如连体防护服可能导致目标不清晰)等因素,可能影响模型的识别准确率和稳定性。需要持续优化模型对恶劣环境的适应能力,并结合多传感器融合等手段提升鲁棒性。

c) 知识库的动态维护与质量。化工行业的安全规程、操作标准、设备信息等知识是动态变化的。如何高效、准确地构建、更新和审核企业安全知识库,确保其时效性和权威性,是保障 RAG 机制有效性的关键,也是一项持续性的工作。

d) 数据安全与合规红线。系统处理的视频、图像数据涉及生产安全和员工隐私,必须严格遵守国家及行业的数据安全法规和标准。数据脱敏、加密传输、访问控制、合规审计等措施是不可或缺的底线要求。

e) “AI幻觉”与信任边界。大模型偶尔会产生不符合事实的“幻觉”输出。在安全攸关的化工领域,如何有效识别并过滤模型的错误判断,建立清晰的人机协同与决策复核机制,明确 AI 的能力边界,是确保系统安全可靠应用的前提。

4.2 发展趋势与展望

面向未来,基于多模态大模型的智能风险识别技术在化工行业的应用将呈现以下发展趋势,并有望深刻变革传统的安全管理模式。

a) 从“识别”到“预测”。结合时序数据分析、工艺参数关联和历史风险数据,模型有望从当前的风险事件识别,向风险趋势预测和早期预警方向发展,例如利用多模态图神经网络进行动态风险预测,实现更主动的安全管理。

b) 领域模型与边缘智能深化。通用大模型将进一步与化工行业知识深度融合,形成更专业的领域大模型。同时,轻量化、高效的边缘侧模型将得到发展,以实现更多风险类型的实时、低成本识别与响应。

c) 人机协同与管理流程重塑。AI 将深度嵌入安全管理流程,从辅助巡检、智能监护,到自动生成安全报告、辅助应急决策、再到提供个性化安全培训。人机协同将成为常态,重塑安全管理工作模式。

d) 与数字孪生、工业互联网融合。与化工厂的数字孪生系统相结合,可在虚拟空间中进行风险模拟、预案推演。与工业互联网平台相融合,实现安全数据与生产、设备数据的联动分析,挖掘更深层次的安全洞察。此外利用对抗训练等策略提升动态场景下的多模态风险预测能力也是未来的研究方向。

综上所述,基于多模态大模型的智能风险识别系统代表了化工安全管理数智化转型的重要方向。尽管面临挑战,但该系统通过知识增强的深度场景理解和推理能力,为提升化工行业本质安全水平、优化管理效率提供了强大的技术引擎。随着技术的不断成熟和应用场景的持续探索,该系统有望成为未来智慧化工安全体系的核心组成部分,助力行业安全管理迈向更智能、更高效、更可靠的新阶段。

参考文献:

- [1] 陶婷婷. 化工园区智能安全监测系统架构研究[J]. 工业安全与环保, 2023, 49(12): 18-27.
- [2] 孟祥龙. 改进 YOLOv7 算法的工业安全多目标检测[J]. 计算机工程与应用, 2024, 60(7): 112-116.
- [3] 吴晓东, 黄俊杰, 刘振宇. 面向化工安全的轻量化 YOLO 动态检测优化[J]. 邮电设计技术, 2025(3): 33-40.
- [4] 应急管理部科技和信息化司. 高危行业 AI 安全监管白皮书(2024 版)[R]. 北京: 中国应急管理出版社, 2024.
- [5] 张立峰, 周子涵, 李航, 等. 工业场景多模态数据融合的实时风险推理框架[J]. 邮电设计技术, 2024(12): 61-64.
- [6] 周子涵, 王雪, 徐天宇. 知识增强的视觉语言模型工业场景适配方法[J]. 自动化学报, 2024, 50(5): 987-998.
- [7] ZHANG Q, LI Y, GUPTA R, et al. VLM-Industrial: Benchmarking Vision-Language Models in Hazardous Scenarios[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(5): 1234-1244.
- [8] CHEN H, LIU B, KIM S, et al. PromptCraft: Structured Prompt Engineering for Industrial VQA[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2024: 889-901.
- [9] SOME L, YANG W, BAIN M. A Comprehensive Survey on Integrating LLMs with Knowledge-Based Methods[J]. IEEE Transactions on Neural Networks and Learning Systems, 2025, 36(2): 567-572.
- [10] 王雪, 李航, 陈明. 工业知识图谱与多模态大模型协同推理研究[J]. 计算机研究与发展, 2025, 62(4): 789-800.
- [11] LIU B, YU C. YOLO-GP: A Multi-Scale Dangerous Behavior Detection Model[J]. IEEE Access, 2024, 12: 45678-45689.
- [12] GUPTA R, ZHANG T, WANG F, et al. Real-Time Anomaly Detection in Petrochemical Plants[J]. IEEE Transactions on Industrial Informatics, 2024, 20(8): 6543-6558.
- [13] LI Y, WANG X, CHEN H, et al. Edge-Cloud Collaborative Video Analytics for Plant Safety[J]. ACM Transactions on Sensor Networks, 2024, 20(3): 1-25.

作者简介:

杨成玮, 工程师, 硕士, 主要从事 5G 工业互联网、智能制造等方面的研究实践工作; 陆新建, 工程师, 硕士, 主要从事化工行业智能化软件方案设计和开发工作; 徐滨阳, 高级工程师, 学士, 主要从事工业互联网、企业数字化咨询和项目管理; 周致永, 工程师, 硕士, 主要从事工业互联网、智能系统设计与优化工作。