

# 사고유형에 따른 유해화학물질 학습데이터 수집 방법에 관한 연구

김연진, 윤아롱, 유성민\*, 심보리, 백성규, 백성하\*\*, 조숙경, 김정배

서원대학교, \*(주)우경정보기술, \*\*인하공업전문대학

anne6497@naver.com, yardkfhd@naver.com, \*smin9830@gmail.com, bori0725@seowon.ac.kr, skbaek@ohcomon.com, \*\*bshzeratul@gmail.com, skyindb@naver.com, gbkim@seowon.ac.kr

## A Study on the Method of Collecting Training Data on Hazardous Chemicals According to Accident Types

Kim Yeon Jin, Yoon A Rong, Ryoo Seong Min\*, Sim Bo Ri, Baek Seong Kyu, Baek Seong

Ha\*\*, Cho Sook Kyoung, Kim Gyoung Bae

Seowon Univ., \*WooKyoung IT, \*\*Inha Technical College

### 요약

본 논문은 유해화학물질 종류를 판독하기 위해 필요한 유해화학물질 학습데이터를 사고유형에 따라 수집하는 방법에 대해 제안한다. 온라인 수집, 소방청 제공영상 수집을 통한 데이터는 영상의 흔들림, 화질 저하 등으로 인해 학습데이터로 사용하기 적합하지 않으므로 본 논문에서는 주로 유해화학물질의 특징과 사고유형을 고려하여 직접 실험을 통해 유해화학물질 학습데이터를 수집하였다. 누출/화재의 사고유형에 따라 학습데이터에서 학습할 수 있는 대상은 불꽃, 연기, 액체이며, 학습대상을 정확하게 어노테이션 할 수 있는 학습데이터를 수집해야 한다. 학습데이터의 다양성을 충족시키기 위한 다양한 조건으로 실험을 진행한 결과, 50만 건 이상의 학습데이터 구축이 가능한 유해화학물질 학습데이터를 수집하였다.

### I. 서론

국내 화학산업이 급성장함에 따라 유해화학물질의 제조·취급·유통량 등이 증가하고 있다. 2020년 화학물질 취급업체 사업장에서는 3만 1,600종의 화학물질이 6억 8,680만 톤 유통되었으며, 유해화학물질 1,413여 종은 총 7,936만 톤이 제조 및 수입되었다[1,2]. 화학물질 유통량 증가로 사용량 또한 증가하면서 다양한 화학사고가 발생하기 시작하였으며, 이는 대부분 작업자 부주의, 화학물질 생산·취급시설의 결함 및 노후와 등으로 인해 발생한다[1].

유해화학물질의 특성상 사고가 발생하면 화학물질 기체 등의 빠른 확산과 분산으로 인해 피해규모가 매우 커지며, 이에 신속한 대비·대응체계를 강화하고자 소방청은 부처합동 전국 7개 화학센터를 운영하고 있다[3]. 그러나 화학사고 재난전문가의 부족으로 올바른 사고물질 판별이 어려우며, 사고물질 방재정보가 책자나 엑셀 파일로 제공되는 등 대응 시스템이 즉각적이지 않아 신속한 대응방안을 찾기에 많은 시간이 소요된다는 문제에 직면하고 있다[4]. 이를 해결하기 위해 소방청은 2022년에 유해화학물질 데이터의 인공지능 학습과 판독 알고리즘을 통해 화학물질을 예측할 수 있는 AI융합 유해화학물질 판독시스템을 구축하였다[5]. 해당 시스템을 통해 도출된 결과를 실제 사고현장에서 활용하기 위해서는 유해화학물질에 대한 방대한 양의 데이터를 활용한 많은 학습이 필요하다. 그러나 실제 유해화학물질 데이터를 새롭게 직접 구축하기에 제한적이므로 현재는 유해화학물질에 대한 데이터가 거의 존재하지 않으며 관련 연구 또한 미비하다.

본 논문에서는 유해화학물질을 탐지하고 판독에 필요한 유해화학물질 학습데이터를 구축하기 위해 사고유형에 따른 유해화학물질 학습데이터 수집 방법에 대해 제안한다.

### II. 본론

#### 1. 유해화학물질 학습데이터 수집

질산, 브롬 등 유해화학물질 14종에 대한 학습데이터 수집은 인터넷 검색을 통한 온라인 수집, 소방청이 보유한 실제 사고현장 출동영상 수집, 실제 실험을 통한 직접 수집으로 총 세 가지 방법을 통해 수집한다. 위험도가 높아 직접 실험하기 어렵거나 실험을 통해 물질의 특성이 나타나지 않았던 물질에 대해서는 온라인 수집을 통해 학습데이터를 확보하였다. 소방청이 보유하고 있는 실제 사고현장 출동영상은 시너, 경유/등유 등 사고가 빈번하게 발생하는 물질에 대한 영상이다. 그러나 인터넷 수집영상과 소방청 제공영상은 영상이 흔들리거나 화질이 현저하게 낮은 영상이 많아 학습용 데이터로 사용하기 적합하지 않았다. 따라서 본 논문에서는 주로 직접 실험을 통해 유해화학물질 학습데이터를 수집하였다.

#### 2. 실험을 통한 유해화학물질 학습데이터 수집

유해화학물질은 물질별 고유의 특징과 색상을 가지며 고체/액체/기체 중 한 가지 상태로 존재한다. 유해화학물질 사고는 크게 누출과 화재의 유형으로 발생하며, 사고유형에 따라 학습데이터에서 학습할 수 있는 대상은 불꽃, 연기, 액체가 된다. 따라서 학습대상을 정확하게 어노테이션 할 수 있는 영상을 수집하고 구축하는 것이 중요하다.

본 논문에서는 학습데이터를 수집하기 위해 주로 유해화학물질 14종에 대한 누출실험, 화재실험을 진행하였으며, 유해화학물질의 잔여 가스 처리가 가능하고 집진 시설이 갖춰진 안전한 장소에서 안전수칙을 준수하며 실험을 진행하였다. 실험 시 물질 영상을 촬영한 장비로는 DSLR 카메라(6대), 갤럭시휴대폰(1대), 열화상카메라(1대)를 사용하였다. [표 1]과 같이 데이터 다양성을 충족하기 위해 농도, 조도, 촬영거리, 촬영각도, 날씨 등










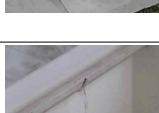
다양한 조건으로 실험을 진행하였다.





[표 1] 유해화학물질 학습데이터의 다양성

조건	내용
농도	진함, 연함으로 구분
조도	영상에서 일부의 조도 범위를 실내/야외촬영
촬영거리	불꽃 및 연기의 크기를 고려하여 거리를 카메라마다 다르게 설정하여 2가지 이상 데이터 수집(약 90~320cm)
촬영각도	동일 물질을 3개 이상의 카메라로 촬영하여 데이터 수집 (각도 수평각: 우측(0°~70°), 정면(71°~110°), 좌측(111°~180°))
날씨	밀폐된 실내공간에서 촬영 (법적 제약 등으로 외부환경에서 촬영이 어려움)

50만건 이상의 학습데이터 구축이 가능하도록 사고유형별로 다양한 실험을 통해 유해화학물질 학습데이터를 수집하였으며, 수집한 학습데이터에 대한 결과는 [표 2]와 같다.

[표 2] 실험을 통한 사고유형별 유해화학물질 학습데이터 수집 결과

실험 물질	실험 데이터	사고 유형	분석
질산		누출/폭발	<ul style="list-style-type: none"> <li>•유독한 연기</li> <li>•공기보다 무거워 바닥에 고임</li> <li>•밝고 트인 환경에서 촬영</li> </ul>
브롬		누출	<ul style="list-style-type: none"> <li>•유독한 연기</li> <li>•질산과 비슷하나 고이기보다는 날림</li> <li>•밝고 트인 환경에서 촬영</li> </ul>
암모니아		누출	<ul style="list-style-type: none"> <li>•무색 투명 액체</li> <li>•특유의 냄새</li> <li>•밝은 환경에서 촬영</li> </ul>
수산화나트륨		누출	<ul style="list-style-type: none"> <li>•가열되고 에탄올이 날아가면 특유의 노란 불꽃</li> <li>•밝은 환경보다는 어두운 환경에서 뚜렷이 관찰됨</li> </ul>
질산칼륨		화재	<ul style="list-style-type: none"> <li>•가열되고 에탄올이 날아가면 특유의 보라색 불꽃</li> <li>•밝은 환경보다는 어두운 환경에서 뚜렷이 관찰됨</li> </ul>
시너		누출/화재	<ul style="list-style-type: none"> <li>•유류계열 화재의 특성이 뚜렷함</li> <li>•밝은 환경에서는 연기도 잘 보이나 어두운 환경에서는 불꽃만 잘 보임</li> </ul>
등유		누출/화재	<ul style="list-style-type: none"> <li>•유류계열 화재의 특성이 뚜렷함</li> <li>•밝은 환경에서는 연기도 잘 보이나 어두운 환경에서는 불꽃만 잘 보임</li> </ul>
염화제2구리		누출	<ul style="list-style-type: none"> <li>•가열되고 에탄올이 날아가면 특유의 초록색 불꽃</li> <li>•밝은 환경보다는 어두운 환경에서 뚜렷이 관찰됨</li> </ul>
염소		누출	<ul style="list-style-type: none"> <li>•노란색의 유독한 연기</li> <li>•어두운 날 실외에서 가장 눈에 띄게 관찰 가능</li> <li>•밝은 장소에서는 색이 옅어지며, 어두운 장소에서는 색은 진하나 잘 보이지 않음</li> </ul>
플루오라이드칼륨		누출	<ul style="list-style-type: none"> <li>•무색 투명 액체</li> <li>•물에 섞은 직후 약간의 발열</li> <li>•약간 어두운 환경에서 관찰이 쉬움</li> </ul>

실험 물질	실험 데이터	사고 유형	분석
황산		누출	<ul style="list-style-type: none"> <li>•무색 투명 액체</li> <li>•약간 어두운 환경에서 관찰이 쉬움</li> </ul>
나이트로벤젠		화재	<ul style="list-style-type: none"> <li>•발화점이 높아 에탄올로 가열하여 실험</li> <li>•화재시 엄청난 양의 검댕이 연기와 함께 날림</li> </ul>
톨루엔		화재	<ul style="list-style-type: none"> <li>•유류계열 화재의 특성이 뚜렷함</li> <li>•밝은 환경에서는 연기도 잘 보이나 어두운 환경에서는 불꽃만 잘 보임</li> </ul>
메틸알코올		누출/화재	<ul style="list-style-type: none"> <li>•연기가 없는 파란색 불꽃</li> </ul>

### III. 결론

본 논문에서는 유해화학물질을 판독하기 위해 필수적인 학습데이터를 구축하기 위한 유해화학물질 학습데이터 수집 방법에 대해 제안하였다. 유해화학물질이 가지는 고유한 특징을 고려하여 유해화학물질 학습데이터를 수집하였고, 학습데이터 수집 방법 중 온라인 수집과 소방청 제공영상 수집을 통한 데이터는 영상의 흔들림, 화질 저하 등으로 인해 학습데이터로 사용하기 적합하지 않았다. 따라서 본 논문에서는 직접 실험을 통해 누출/화재의 사고유형에 따른 유해화학물질 학습데이터를 수집하였으며, 학습데이터의 다양성을 만족하기 위해 조도, 촬영거리, 촬영각도 등을 고려하여 여러 가지 조건으로 실험한 결과, 50만 건 이상의 학습데이터 구축이 가능하도록 유해화학물질 학습데이터를 수집하였다. 사고유형에 따라 화재사고에서는 물질별로 나타나는 고유한 화재 특성이나 불꽃색을 보였으며, 누출사고에서는 물질별 상이한 기체 확산 특징이나 연기색을 나타냈다. 향후 환경의 변화로 인해 불꽃, 연기의 관찰이 불가능하고, 무색/무취로 시각적 특징이 없는 유해화학물질의 판독까지 가능할 수 있도록 분광 스펙트럼 데이터 및 열화상 데이터 등을 활용한 유해화학물질 학습데이터 생성 방법에 관한 연구가 지속적으로 진행되어야 한다.

### ACKNOWLEDGMENT

본 연구는 과학기술정보통신부와 정보통신산업진흥원 주관으로 소방청 컨소시엄에서 수행하는 “AI융합 유해화학물질 판독시스템 사업” (2022~2024)의 지원을 받았다

### 참고 문헌

- [1] 이선자, “화학사고 예방을 위한 안전관리 선진화에 관한 연구”, 박사학위논문, 명지대학교, 2020.
- [2] 환경부, 제4차 화학물질 통계조사 결과, (<http://www.me.go.kr/>)
- [3] 이태형, 이덕재, 신창현, “화학사고에 의한 인명사고 특성 분석에 관한 연구”, 한국화재소방학회논문지, Vol. 31, No. 1, pp.81-88, 2017.
- [4] 조숙경, 백성하, 박봉섭, 김경배, “인공지능 기술을 활용한 유해화학물질 탐지 방법”, 한국통신학회 추계종합학술발표회 논문집, pp.264-265, 2022.
- [5] 김연진, 조숙경, 박봉섭, 김경배, “유해화학물질 판독을 위한 인공지능 학습 데이터 라벨링에 관한 연구”, 한국통신학회 추계종합학술발표회 논문집, pp.266-267, 2022.