

# AI 프레임워크를 활용한 작업 문서 자연어 기반 멀티모달 이상감지 모델 적용의 관한 연구

\*박성일, 김정식

\*엔소프트, 한국전자통신연구원

\*sipark@nsoft.co.kr, j.s.kim@etri.re.kr

## Research on the application of natural language-based multimodal anomaly detection models for working documentation using an AI framework.

\*Park Sung Il, Jeongsik Kim

\*Nsoft Corp., ETRI

### 요 약

본 연구는 AI 프레임워크를 활용하여 작업 문서의 자연어와 멀티모달 데이터를 기반으로 한 이상감지 모델을 개발하고 적용하는 것을 목적으로 한다. 현대 산업 환경에서 작업 문서의 정확성과 일관성은 생산성과 안전성에 직결되는 중요한 요소이다. 이에 본 연구에서는 텍스트를 기반으로 하는 문서에, 환경 정보, 기상 정보 등 다양한 형태의 메타 데이터를 통합적으로 분석하여 잠재적인 이상을 감지하는 AI 모델을 제안한다.

### I. 서 론

본 논문에서는 기간계 시스템에서 발생하는 다양한 데이터 중 문서를 기반으로 하는 자연어 데이터를 주로 사용하여, 사용자 안전을 위해하는 위험요인을 탐색하고 찾아내는 멀티모달 모델을 제안한다.

연구 방법으로는 먼저 대규모 작업 문서 데이터셋을 수집하였다. 한국산업안전보건공단에서 제공하는 국내 재해사례 리포트를 수집하여 AI 모델을 위한 전처리를 수행하였다. 또한 이미지 데이터, 환경 정보 데이터 등을 처리하기 위한 sLLM 모델을 접목하고, 최종적으로 이상 감지를 위해 오토인코더 기반 알고리즘을 적용하였다.

이러한 실험을 통해 단일 모빌리티 기반 모델의 한계를 벗어나, 멀티 모달 데이터(비전, 시계열, 자연어)를 기반으로 하는 이상감지 모델을 구성할 수 있었다. 이러한 모델은 다양한 데이터의 결합과 순서 정보의 반영으로 보다 복잡한 분석을 가능하게 한다.

본 연구의 결과는 산업 현장에서의 작업 문서 관리 및 품질 제어 프로세스, 작업자 작업 환경 안전 개선에 직접적으로 적용될 수 있을 것으로 기대된다. 향후 연구에서는 실시간 처리 능력 향상과 보다 다양한 데이터를 결합하는 멀티모달 데이터 구축, 그리고 최종적으로 산업현장의 작업환경 전체를 종합적으로 판단하는 AI 인지 모델을 제시하는 것을 목표로 한다.

### II. 본 론

본 논문에서 해결하고자 하는 산업 현장 작업자의 안전을 위한 멀티모달 데이터 기반 이상감지 모델은, 단일 모빌리티 기반의 이상감지 모델의 한계를 넘어 멀티 모달 데이터를 활용한 보다 복잡한 이상감지 구조를 구현하여 산업 현장 작업의 위험성을 다면적으로 이해하는 모델을 구성하기 위함이다.

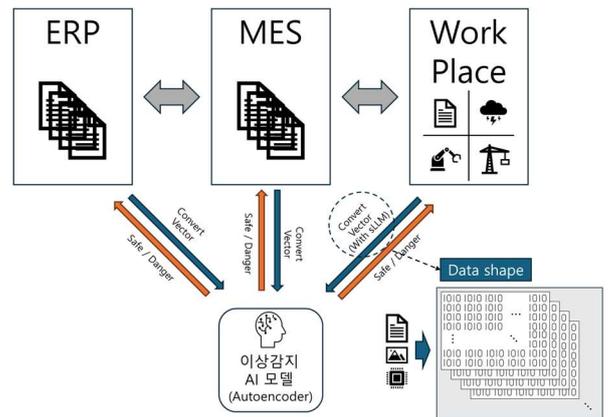


Fig 1. Connecting multimodal data from industrial sites with anomaly detection models.

일반적인 산업 환경에서 “제품”을 제조하는 전 단계는 다양한 기간계 시스템을 복합적으로 사용하여 전 과정을 진행한다. 이러한 시스템에서 발생하는 데이터는 대표적으로 문서(생산계획서, 작업지시서 등)가 있으며 이에 포함되는 메타데이터또한 시스템에서 발생하는 대표적인 데이터이다.

기존의 이상감지 모델은 단일 모빌리티 기반(예를 들면 신호처리, 진동 등)의 이상감지로 단일 영역에서의 비정상성을 판단하는 모델은 이미 다양한 산업 분야에서 활용되고 있다. 대표적으로 각종 센서, PLC, DCS, SCADA 등을 통해 수집된 데이터를 분석하여 비정상성을 판단하는 모델은 제조사 레벨, 혹은 그 상위 구성(제어부, 플랫폼 등)에서 이미 보편적으로 활용 중이다.

본 연구에서는 기본적인 이상감지를 수행할 대표 모델로 오토인코더 기반의 모델을 구성하였다. 이는 오토인코더의 특성에 기인한 것으로, 입력 데이터의 특징을 추출하는 이 신경망 구조가 본 논문에서 제시하는 이상감지 구조에 적합한 것으로 사료되었다.

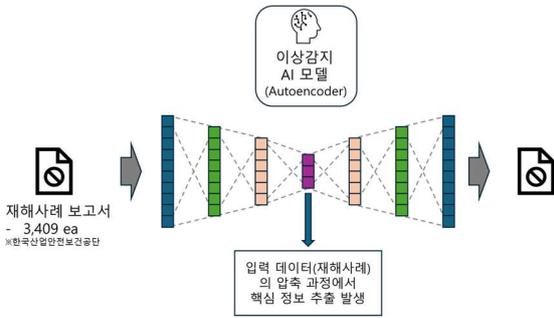


Fig 2. Information compression for disaster case feature extraction.

위와 같이 비정상 상황(재해 사례)의 데이터의 활용을 통해 오토인코더 모델은 재해 상황의 특징을 추출하고 학습한다. 이후 이 모델에 정상 상황의 데이터를 입력하면 정상 복원하지 못하여 오차가 커질 것이며, 학습한 데이터와 비슷한 데이터(재해 사례)가 입력되면 정상적으로 복원할 수 있을 것이다.

이렇게 구성된 오토인코더 모델에는 단일 문서(예를들면 작업지시서)만이 입력값으로 사용되는 것이 아닌, 다양한 추가적인 정보를 포함하여 활용할 수 있다. 이 부분이 멀티모달 데이터의 핵심적인 부분이며, 이후 이상감지 모델 자체를 개선하는 부분에도 크게 기여할 수 있는 영역이다.

지금까지 본 연구에서는 일반적으로 생각할 수 있는 “문서” 형태의 데이터를 설명하고 있다. 이는 지류 문서뿐만 아니라, 시스템에서 발생하는 전자 문서 또한 포함된다. 이러한 문서(자연어) 데이터를 전처리(토큰화, 임베딩 등)하여 모델이 이해할 수 있는 형태로 변환한 뒤 활용하였다. 하지만 산업 현장에서 발생하는 다양한 비정형 데이터를 기존의 데이터와 결합하여 제시한 모델에 활용 및 접목하고자 한다.

산업 현장에서는 기본적으로 PLC, DCS, SCADA 와 같은 센서 혹은 장비에서 발생하는 데이터에 더해, 최근에는 영상 데이터 또한 다양하게 수집하고 활용되고 있다. 이러한 비정형 데이터는 최근 주목도가 높은 LLM(혹은 sLLM)의 모델을 활용하여 자연어 형태의 데이터로 변환할 수 있다.



Fig 3. Transformation process of unstructured data (images) with LLM and application of anomaly detection model.

Fig3 은 그중 하나의 예시로서 제시하고자 한다. 산업 현장에서 수집된 비정형(이미지) 데이터는 LLM 모델을 통해 사용자 행위, 객체 정보를 설명하는 자연어 데이터로 변환된다. 이후 자연어 전처리 과정을 통해 해당 내용은 벡터로 변환된다. 그리고 최종적으로 이상감지 모델에 입력되어 정상/비정상 상태를 판별하게 된다.

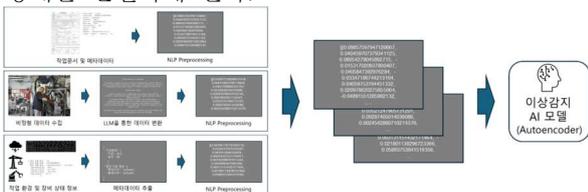


Fig 4. Applying anomaly detection autoencoders with multimodal data (documents, images, sensors, etc.).

본 연구에서는 Fig4 에서 제시한 바와 같이 산업 현장에서 발생하는 멀티모달 데이터를 벡터로 변환하여 이상감지 모델에 적용하였다. 재해사례 데이터를 학습한 오토인코더 모델은 입력받은 멀티모달 기반 데이터를 받아 정상/비정상 여부를 판별하고 그 결과를 제시한다.

실제 AI 프레임워크를 통해 구축한 이상감지 서비스는 백엔드 기반으로 작동하나, 시각적 이해를 위한 간단한 웹 서비스를 구성하였다(Fig5). 이와 같은 간단한 웹서비스 구조를 참고하여 타 서비스와의 연계 및 통합할 수 있으며, 이 웹 서비스 내부 구조는 실제 AI 프레임워크 내부에서 restful API 를 구성하여 연계하는 구조와 동일하게 하였다.

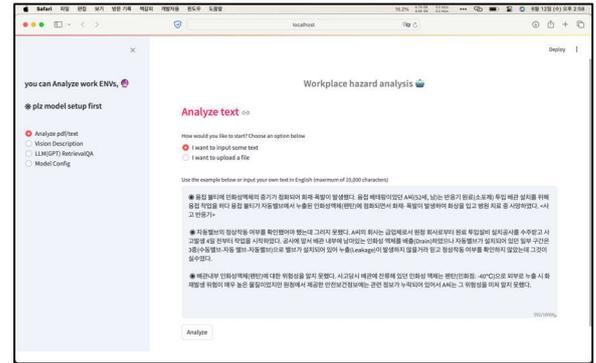


Fig 5 Configuring web services for multimodal anomaly detection autoencoders

### III. 결론

본 논문에서는 작업자의 안전을 위한 멀티모달 기반 이상감지 모델의 구성과 적용 방법에 대해 연구하였다. AI 프레임워크를 활용하여 실제 산업 환경에서 멀티모달 이상감지 모델을 테스트하였고, 그 과정에서 다양한 데이터(문서, 시스템 로그, 센서 등)를 통합하는 것으로 복합적인 위험요인 판별하는 방법을 제시하였다. 그러나 이 모델을 보다 넓은 산업 영역에 접목하는 것, 다양한 추가 데이터를 통합하는 것 그리고 일관된 성능을 보장하는 것은 추가적인 향후 연구 과제로 사료된다.

### ACKNOWLEDGMENT

본 논문은 한국전자통신연구원 기본사업과 울산광역시-ETRI 공동협력사업의 지원을 받아 수행되었음. [24ZB1200, 인간중심의 자율지능시스템 원천기술 연구, 24AS1600, 제조 혁신을 위한 주력산업 지능화 기술 개발 및 산업현장에서의 사람-이동체-공간 자율협업지능 기술 개발]

### 참 고 문 헌

- [1] Zhao Minglu, Furuha Reo, Agung Mulya, Takizawa Hiroyuki, Soma Tomoya “Failure Prediction in Datacenters Using Unsupervised Multimodal Anomaly Detection”, IEEE International Conference on Big Data, pp.3545-3549(2020).
- [2] Zhao Yingying, Liu Qi, Li Dongsheng, Kang Dahai, Lv Qin, Shang Li “Hierarchical Anomaly Detection and Multimodal Classification in Large-Scale Photovoltaic Systems”, IEEE transactions on sustainable energy, pp. 1351-1361(2019-07)