

양상블 인공지능 모델을 활용한 안전 관리 자율운영 플랫폼 설계

Safety Autonomous Platform Design with Ensemble AI Models

Dongyeop Lee^{1,2*} · Daesik Lim^{1,2} · Soojeong Woo¹ ·
Youngho Moon¹ · Minjeong Kim¹ · Joonwon Lee²

¹Convergence Technology Laboratory, KEPCO Research Institute, Daejeon, 34056, Korea
²Department of Safety Health Convergence Engineering, Soongsil University, Seoul, 06978, Korea

[Abstract]

This paper proposes a novel safety autonomous platform (SAP) architecture that can automatically and precisely manage on-site safety through ensemble artificial intelligence models generated from video information, worker's biometric information, and the safety rule to estimate the risk index. We practically designed the proposed SAP architecture by the Hadoop ecosystem with Kafka/NiFi, Spark/Hive, Hue, ELK (Elasticsearch, Logstash, Kibana), Ansible, etc., and confirmed that it worked well with safety mobility gateways for providing various safety applications.

Key word : Autonomous platform, Ensemble AI, Safety management, Power industry, Safety rule.

1. 서론

ISO45001는 산업 보건 및 안전관리 경영시스템의 국제 표준으로 전 세계 사업장에서 작업 현장의 위험을 줄이고 작업자의 안전을 향상시키며 안전한 근무 조건을 조성하기 위한 프레임

워크를 제공하고 직원의 부담을 줄이기 위한 요구사항 등을 포함하고 있다[1]. 이에 전력산업과 관련된 다양한 작업 현장에서도 이러한 ISO45001에 기반한 중대재해처벌법에 대응하기 위해 첨단화된 안전관리체계를 마련할 필요가 있다[2]. 특히 전력산업 현장은 약 230종의 위험성 체크 리스트를 중심으로 공종별 안전관리 업무를 체계화하여 관리해야 하는 어려움이 있다.

따라서 본 논문에서는 전력산업 현장에 특화된 안전관리 체계에 따라 GIS(geographic information system)의 현장 영상정보와 작업자의 생체정보(BI; biometric information)를 수집하고, 작업자의 위해 요인을 판단 및 예측할 수 있는 양상블 인공지능(AI; artificial intelligence) 모델을 통해 위험도 지수(RI; risk index)를 추정하는 안전관리 자율운영 플랫폼을 제안한다.

II. 안전관리 자율운영 플랫폼 설계

그림 1은 제안한 안전관리 자율운영 플랫폼(SAP; safety autonomous platform) 아키텍처를 나타낸다. 이 플랫폼은 크게 안전 데이터 관리 모듈, AI기반 안전판단 메인엔진, GIS기반 작업자 안전관리 시각화로 구성되어 있다.

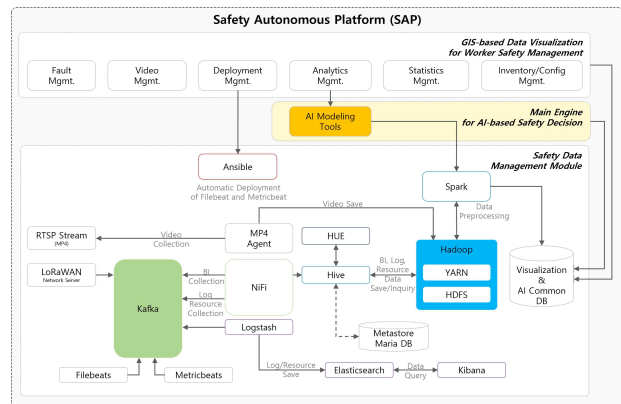


그림 1. 안전관리 자율운영 플랫폼 아키텍처
Fig. 1. Safety autonomous platform (SAP) architecture.

먼저 플랫폼의 근간이 되는 안전 데이터 관리 모듈은 하둡 에코시스템으로 설계되었다. 일반적으로 빅데이터 플랫폼은 수집과

<https://doi.org/10.12673/jant.2024.28.1.159>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 9 January 2024; Revised 26 February 2024
Accepted (Publication) 28 February 2024 (29 February 2024)

*Corresponding Author; Dong-Yeop Lee

Tel: +82-42-865-5429

E-mail: iamdylee@kepco.co.kr

정제 단계, 분석과 시각화 단계를 통해 처리되며, 이러한 단계별 처리과정 기술이 완전한 선순환 사이클로 구성될 때 빅데이터 기반 플랫폼이라고 한다. 하둡 코어 프로젝트는 분산 데이터 저장을 위한 HDFS (hadoop distributed file system)와 자원 관리를 위한 YARN (yet another resource negotiator)으로 구성되었고, SQL (structured query language)로 분산 데이터를 처리하기 위해 Hive와 인-메모리 기반의 고속 데이터 처리를 위해 Spark를 서브 프로젝트로 사용하였다[3]. 특히 Spark는 머신러닝 기반 배치 처리, SQL 기반 질의처리, 실시간 스트리밍 데이터 흐름 처리, GUI 라이브러리 처리와 같은 다양한 작업을 수용할 수 있다. 또한 웹 기반 사용자 인터페이스로 HUE (hadoop user experience)를 사용하여 다른 하둡 에코시스템과 함께 그룹화되어 Hive 작업과 Spark Job 등의 실행이 가능하도록 하였다[4]. NiFi는 실시간 데이터 처리흐름을 모니터링하기 위한 오픈소스 프레임워크로써, 다양한 노드를 통과하는 데이터 흐름을 Web UX에서 그래프로 표현하며 규격과 포맷 형식이 다르더라도 분석이 용이하고 데이터를 처리할 때 제어상태에 따른 우선 순위 부여가 가능하다. ELK (ElasticSearch, Logstash, Kibana) 스택은 데이터를 인덱싱하고 분석하고 검색하는 ElasticSearch, 데이터를 수집 및 변환하여 특정 대상에게 전송하는 Logstash, 데이터 분석 결과를 시각화하는 Kibana로 구성된 스택을 말하며, 다양한 시스템과 응용에서 로그를 수집 및 합산하고 이를 통해 응용과 인프라 모니터링 시각화 UX를 생성하여, 발생한 문제를 빠르게 파악하고 해결할 수 있다는 장점이 있다[5]. Kafka는 실시간 데이터 스트림을 수집 및 관리하기 위한 멀티 환경 지원 분산 시스템으로 퍼블리싱 (publish)-구독(subscribe) 모델로 구성되어 있으며 시스템 안전성을 위해 로드밸런싱과 내고장성을 보장하고 있다. 전력산업 현장과 작업자들에 의해 발생하는 실시간 데이터를 수집하기 위해 이러한 Kafka를 사용하였다.

AI기반 안전판단 메인엔진은 안전 데이터 관리 모듈과의 연동을 통해 작업 현장 CCTV와 작업자 안전모 카메라의 영상정보(데이터), 작업자의 생체정보(데이터)를 기반으로 작업 현장과 작업자에게 적합한 안전판단 이상불 인공지능 모델을 생성하고 배포하는 기능을 수행하고, 이 이상불 인공지능 모델의 추론 결과, 전력산업 현장의 표준 작업 절차 (SOPs; standard operating procedures), 작업자의.프로파일 등을 통해 위험도 지수를 추정하는 역할을 수행한다. GIS기반 작업자안전관리 시각화는 전력산업 현장의 공간 정보를 기반으로 작업 현장과 작업자의 안전관리와 안전상태를 종합적으로 모니터링하기 위한 통합 대시보드이며 위험도 지수를 고려하여 현장의 실시간 공중별 현황을 레이어별로 구축하고 클러스터링하였다.

그림 2는 안전관리 자율운영 플랫폼과 전력산업 현장에 위치한 안전관리 모니터링 게이트웨이(SMG; safety mobility gateway)와의 연동 과정을 도식화한 것이다. 안전관리 모니터링 게이트웨이는 작업자의 안전모에 장착된 카메라 센서와 작업현장 CCTV로부터의 영상 정보를 RTSP (real-time streaming protocol) 서버로 수집하여 안전관리 자율운영 플랫폼으로 전달하고, 작업자의 생체정보(BI)도 LoRa Proxy로 수집하여 전달한다. 또한 Filebeats를 통해 사용자가 지정한 로그 파일이나 위치를 모니터링하고 로그 이벤트를 수집하여 전달하고, Metricbeats를 통해 시스템과 서비스의 메트릭 정보(CPU, 메모리, 네트워크 등 모든

프로세스에 대한 통계 정보)를 수집하여 전달한다. 안전관리 모니터링 게이트웨이는 AI기반 안전판단 메인엔진을 통해 생성 및 배포된 이상불 인공지능 엔진을 탑재하고 있으며, 이를 통해 현장 안전관리자는 작업 현장의 공중별 안전상황을 쉽게 파악하고 위험 경고 메시지가 발생할 경우 신속하게 대응할 수 있다.

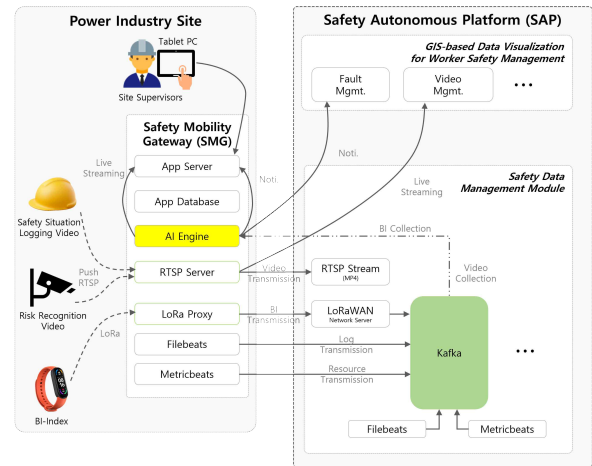


그림 2. 안전관리 자율운영 플랫폼과 안전관리 모바일리티 게이트웨이의 연동
Fig. 2. Interworking with the SAP and the SMG.

III. 이상불 인공지능 모델

3-1 학습데이터셋 구축 절차

AI기반 안전판단 메인엔진에서 작업현장과 작업자 맞춤형 이상불 인공지능 모델을 생성하기 위해서는 체계적인 학습데이터 구축 방안이 요구되며 이를 위해 그림 3과 같은 데이터 정제 및 라벨링 절차를 설계하였다.

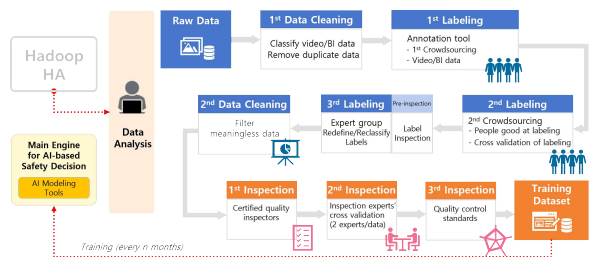


그림 3. 학습데이터 구축을 위한 데이터 정제 및 라벨링 절차
Fig. 3. Data cleaning and labeling procedure for building a training dataset.

하둡 에코시스템을 통해 수집된 원시데이터를 1차 정제하면서 주로 비정형 데이터인 영상정보와 생체정보를 분류하고 중복된 데이터를 제거한다. 정제된 데이터에 대해 어노테이션(annotation) 틀을 사용하여 크라우드소싱(crowdsourcing)으로 1차 레이블링을 수행한다. 다음으로 레이블링을 잘하는 일반인 그룹의 크라우드소싱으

로 2차 레이블링을 수행하는데 기존 1차 레이블링 결과와 교차검증을 통해 레이블링 결과를 검수하고 수정하는 작업을 주로 수행한다. 3차 레이블링 작업은 전문가 그룹을 통해 목적에 부합하는 레이블을 재정의 및 재분류하여 레이블 사진 검수를 진행한다. 이후 2차 데이터의 정제를 수행하면서 무의미한 데이터를 필터링하게 된다. 이렇게 완성된 학습데이터셋은 품질관리 기준에 부합할 수 있도록 3차에 걸친 검수 과정을 거친다. 최종적으로 학습데이터셋이 완성되면 AI기반 안전판단 메인엔진에서 양상블 인공지능 모델을 생성하는데 사용된다. 앞서 언급한 것처럼 생성된 모델은 표준 통신 프로토콜에 따라 도커 컨테이너(docker container) 기반으로 안전관리 모바일리티 게이트웨이로 배포되고 관리된다.

3-2 양상블 인공지능 모델 생성 방법

작업현장과 작업자의 안전 상황을 판단하기 위해 그림 4와 같이 양상블 인공지능 모델을 생성한다. 양상블 인공지능 모델은 여러 개의 개별 인공지능 모델을 학습시켜 그 모델들의 결과들을 조합하여 추론하는 모델로써 개별 모델의 성능을 분산시켜 과적합을 감소시키고 개별 모델의 성능이 나쁠 때 성능을 보강할 수 있는 장점이 있다.

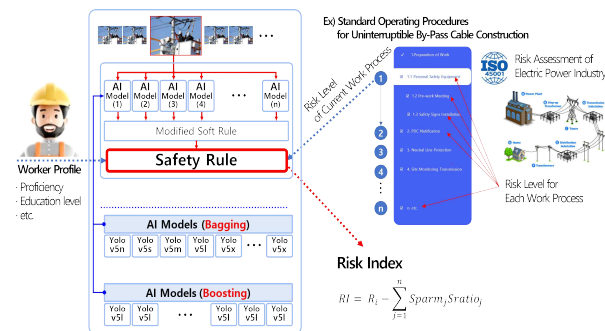


그림 4. 양상블 인공지능 모델의 생성 방법
Fig. 4. Generation method of the ensemble AI models.

일반적으로 양상블 모델은 학습 방식에 따라 배깅(bagging), 부스팅(boosting), 스택킹(stacking)으로 나뉘는데 우리는 안전 상황 판단에 적합한 배깅과 부스팅 방식을 채택하였다. 배깅 방식은 병렬 학습으로 과적합에 강한 장점이 있고 부스팅은 순차 학습으로 정확도가 높다는 장점이 있기 때문에 우리는 작업현장과 작업자의 상황에 따라 두가지 학습 방식을 비교하여 더 나은 학습 방식을 선택적으로 적용하도록 하였다. 또한 양상블 모델의 최종 추론값을 결정짓는 보팅(voting) 방식으로 하드 보팅과 소프트 보팅이 있는데 우리는 변형된 소프트 보팅 방식(modified soft rule)을 사용한다.

표 1. 안전관리 규칙의 파라미터
Table 1. Parameters for the safety rule.

S_{parm_j}	worker proficiency	worker edu. level	work environment	...	etc.
S_{ratio_j}	$Ratio_1$	$Ratio_2$	$Ratio_3$...	$Ratio_n$

양상블 인공지능 모델의 변형된 소프트 보팅 방식에 의해 얻은 최종 추론값은 전력산업 현장의 표준 작업 절차, 작업자의 프로파일 등과 함께 표 1에 나타난 안전관리 규칙(safety rule)의 파라미터로 사용된다. 각 파라미터마다 상황에 따라 정의된 특정한 가중치 비율(ratio)을 갖게 되는데 그 총합은 식(1)과 같이 1이 되게 정의하였다.

$$\sum_{j=1}^n Ratio_j = 1. \tag{1}$$

우리는 안전관리 규칙을 식 (2)와 같이 정의하였으며 안전관리 규칙 파라미터를 입력으로 받아 위험도 지수를 추정하게 된다.

$$RI = R_i - \sum_{j=1}^n S_{parm_j} S_{ratio_j} \tag{2}$$

여기서, R_i 는 i 번째 현장의 표준 작업 절차의 위험성 평가를 통해 얻은 정규화된 위험도를 나타내고, S_{parm_j} 는 위험도를 낮출 수 있는 j 번째 안전관리 파라미터를 나타내고, S_{ratio_j} 는 j 번째 안전관리 파라미터에 대한 가중치 비율을 나타낸다.

우리는 학습 데이터의 종류가 영상정보인지 생체정보인지에 따라 그림 5와 같이 두 종류의 양상블 인공지능 모델을 생성한다. 영상정보일 경우 YOLOv5 모델[6]을 조합하여 양상블 인공지능 모델을 생성하고, 생체정보일 경우 시계열 데이터의 이상 탐지 성능이 우수하다고 알려진 TadGAN(time series anomaly detection generative adversarial network) 모델을 조합하여 양상블 인공지능 모델을 생성한다. 이러한 양상블 인공지능 모델은 여러 지표(mean average precision, anomaly error score 등)에 따른 성능 평가를 거친 후 최적의 모델로 선정되고 안전관리 모바일리티 게이트웨이에 탑재된다. 식(2)를 통해 추정된 위험도 지수는 메시지 큐에 저장되고 경고 알림이나 현장 감독용 응용서비스 등에 활용되게 된다.

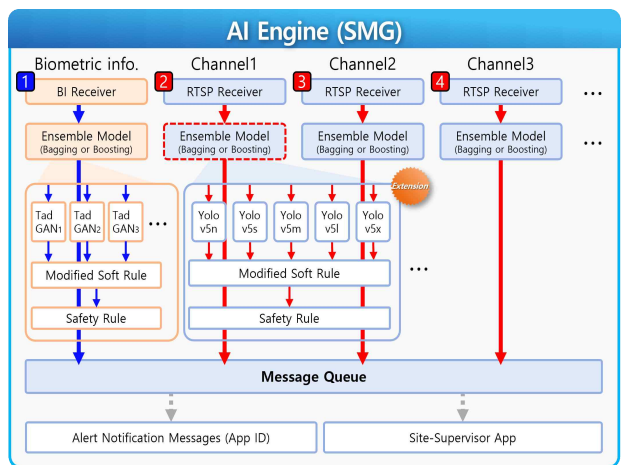


그림 5. 안전관리 모바일리티 게이트웨이에서의 양상블 인공지능 모델의 응용서비스

Fig. 5. Applications of ensemble AI models in SMGs.

IV. 구현 및 테스트 결과

제한한 안전관리 자율운영 플랫폼과 안전관리 모빌리티 게이트웨이 프로토타입을 구현하여 기능성 평가를 수행하였다. 그림 6은 학습 데이터셋 중 영상정보로 생성한 앙상블 인공지능 모델의 테스트 성능평가 예시를 보여주고 있다. 향후 다양한 전력 산업 현장 상황에서 영상정보와 생체정보를 기반으로 앙상블 인공지능 모델을 생성하고 평가하여 인공지능 범용 DB를 구축할 예정이다. 그림 7은 제한한 안전관리 자율운영 플랫폼으로 구현한 응용서비스 예시를 나타낸다. 왼쪽에 보이는 전력 산업 현장의 표준 작업절차에 따라 안전관리 모빌리티 게이트웨이로부터 실시간으로 취득한 영상정보와 생체정보를 모니터링할 수 있으며 앙상블 인공지능 모델과 안전관리 규칙에 따라 추정된 위험도 지수에 의해 장비 주의 알람 2건과 작업자 건강 주의 알람 1건이 수신된 것을 확인할 수 있었다.



그림 6. 앙상블 인공지능 모델의 성능평가
Fig. 6. Performance evaluation of ensemble AI models.

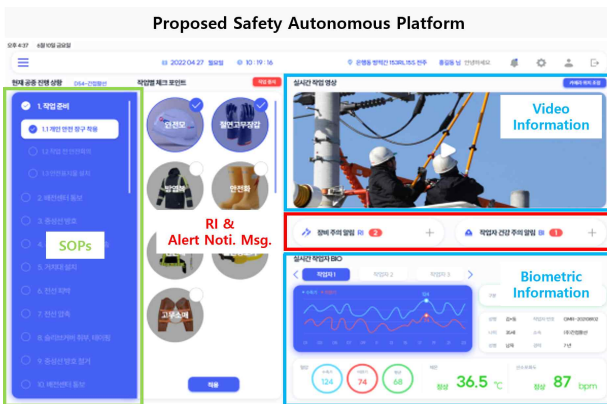


그림 7. 안전관리 자율운영 플랫폼의 응용서비스
Fig. 7. Application for the safety autonomous platform.

V. 결론

본 논문에서는 전력 산업 현장에서 취득한 영상정보와 작업자의

생체정보를 학습 데이터로 생성한 앙상블 인공지능 모델과 그 앙상블 모델의 추론 결과와 작업 현장의 표준 운영 절차에 따른 위험도, 작업자의 프로필 등을 파라미터로 위험도 지수를 추정하는 안전관리 규칙에 따라 작업 현장의 안전을 자동적이고 정확하게 관리할 수 있는 안전관리 자율운영 플랫폼 아키텍처를 제안하였다. 제안한 아키텍처는 크게 하둡 에코시스템 기반의 안전 데이터 관리 모듈, AI기반 안전판단 메인엔진, GIS기반 작업자 안전관리 시각화로 구성되었으며 작업 현장의 안전관리 모빌리티 게이트웨이와 연동하도록 설계되었다. 구현 및 기능성 평가를 통해 앙상블 인공지능 모델의 추론 결과와 안전관리 규칙에 따라 안전관리 자율운영 플랫폼의 주요 기능이 성공적으로 동작하는 것을 확인하였다.

Acknowledgments

본 논문은 한국전력공사의 주력연구개발과제의 지원을 받아 수행된 연구 결과입니다[과제명: 위험성 평가기반 공사현장 자율 안전솔루션 개발/과제번호 : R221A05].

References

- [1] M. Soltanifar, *ISO 45001 Implementation: How to Become an Occupational Health and Safety Champion*, New York, NY: Productivity Press, 2022.
- [2] H. L. Floyd II, "Reflections on a 50-year journey in electrical safety: further steps toward risk reduction," *IEEE Industry Applications Magazine*, Vol. 29, No. 4, pp. 59-65, April 2023.
- [3] X. Sun, Y. He, D. Wu and J. Z. Huang, "Survey of distributed computing frameworks for supporting big data analysis," *Big Data Mining and Analytics*, Vol. 6, No. 2, pp. 154-169, June 2023.
- [4] S. P. R. Asaithambi, R. Venkatraman, and S. Venkatraman, "A thematic travel recommendation system using an augmented big data analytical model," *Technologies*, Vol. 11, No. 1, pp. 1-28, Jan. 2023.
- [5] G. Calderon et al., "Monitoring framework for the performance evaluation of an IoT platform with Elasticsearch and Apache Kafka," *Information Systems Frontiers*, pp. 1-17, July 2023.
- [6] M. Hussain, "YOLO-v5 variant selection algorithm coupled with representative augmentations for modelling production-based variance in automated lightweight pallet racking inspection," *Big Data and Cognitive Computing*, Vol. 7, No. 2, pp. 1-15, June 2023.
- [7] J. Lee, S. Cho, and D. Yu, "Comparative study of time series anomaly detection based on generative deep learning models," *The Korean Data & Information Science Society*, Vol. 34, No. 3, pp. 377-394, May 2023.