

의료 빅데이터기반 확장 인공지능 통합플랫폼

이충섭¹, 김지언¹, 노시형¹, 김태훈^{1,2}, 이윤오², 유영주¹, 천정범³, 정창원^{1,2}

¹원광대학교 의료융합연구센터

²원광대학교병원

³스마트케어웍스

e-mail : {cslee99, medikim3551, nosij123, tae_hoonkim}@wku.ac.kr,
wonduke@empal.com, yerie1.ryu@gmail.com, jb.chun@smartcareworks.com,
mediblue@wku.ac.kr

Medical bigdata-based Extended Artificial Intelligence Integration Platform

Chung-sub Lee¹, Ji-Eon Kim¹, Si-Hyeong Noh¹, Tae-Hoon Kim^{1,2}
Yun Oh Lee², Yeong-Ju Yu¹, JungBum Chun³, Chang-Won Jeong^{1,2}

¹Medical Convergence Research Center, Wonkwang University

²Smart Health IT Center, Wonkwang University Hospital

³SmartCareworks Inc.

요약

최근 의료데이터의 표준화를 기반으로 다양한 임상연구가 국내외에서 활발하게 진행되고 있다. 그러나 대부분 개발기술이 임상현장에 적용되지 못하는 이유는 상이한 인프라로 인한 일관성있는 결가를 도출하지 못하는 문제점과 부족한 진단지표와 기준 그리고 충분하지 못한 기술적·임상적 검증이 문제 가 되고 있다. 본 논문에서는 이러한 문제점을 해결하기 위한 새로운 통합 플랫폼을 제안하고자 한다. 이를 위해서 임상데이터는 OHDSI의 OMOP-CDM으로 표준화되어야 하며, 이외에 의료영상 정보를 포함한다. 제안한 플랫폼은 표준화된 데이터를 통해 지속적인 자가 학습을 수행하며, 질환별 진단에 필요한 개발 도구와 분석 소프트웨어 도구를 통해 다양한 타겟 질환연구를 지원한다. 제안한 플랫폼은 질환에 대한 비침습적 진단을 위해 의료영상을 기반으로 데이터표준화를 기반으로 하며, 이를 통해 인공지능 기술을 개발하고 병원 정보시스템과 연계하여 임상현장에 실증을 통해 검증하고자 한다.

1. 서론

최근 인공지능 산업 분야에서 인공지능 플랫폼의 경쟁으로 진행되고 있다. 이러한 인공지능 플랫폼은 헬스케어 분야에 적용되면서 다양하고 방대한 양의 데이터를 인공지능 기술을 활용하여 학습·분석하고 이를 기반으로 질병 예측, 예방의학, 환자 맞춤형 치료 등 새로운 의료 서비스 제공하는데 핵심으로 급부상하고 있다[1]. 특히, 보건의료 데이터의 폭발적 증가는 기하급수적으로 증가하는 의료 데이터를 신속히 분석하여 전문 의료지식을 도출하고, 이를 의료 서비스의 질적 수준향상과 차세대 의료산업으로 연결시키기 위한 노력을 하고 있다.

인공지능은 보건의료 빅데이터의 분석과 활용에 중점을 두고 있다. 인공지능 기술개발을 위해서는 보건의료 빅데이터가 학습 인프라가 되면서 선순환 구조를 형성하여 관련된 기술개발을 앞당길 수 있다[2].

그러나 기관별 상이한 인프라와 이를 통해 생성되는 데이터는 인공지능 기술 도입을 지원시키는 요소로 표준화에 맞추어 정형화 시켜야하는 이슈가 대두되었다.

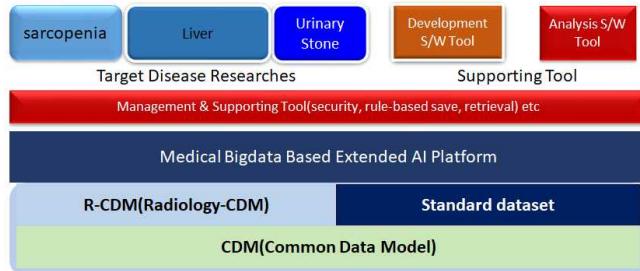
최근 3년 동안 국내에서 정부과제 주도하에 선행공통 데이터모델기반 분산형 바이오헬스 통합데이터망 구축기술개발사업을 진행하고 있으며, 국내 주요 40여개 의료기

관과 함께 해당 병원 내원 환자 5000만명의 의료 데이터 표준화를 목표로 OMOP-CDM으로 변환하고 있다. 그럼에도 불구하고 비정형 임상 데이터의 표준화가 절실하게 요구되고 있다. 특히, 의료영상의 경우 의료영상기반의 질환 진단 및 분석분야에 인공지능 기술을 도입하기 위해서는 학습용 데이터셋 구축이 필요하다. 본 연구는 질환의 비침습적 진단방법으로 의료영상 분석을 포함하기 때문에 의료영상 데이터 수집을 통해 의료영상까지 포함하여 표준화 작업을 수행하였다. 또한, 지속적인 학습을 위한 플랫폼을 설계하고 기존 인프라와의 연계를 통해 질환 진단을 위한 머신러닝 알고리즘 개발이 용이하게 개발환경을 구축하였다.

2. 의료빅데이터기반 확장 인공지능 통합 플랫폼

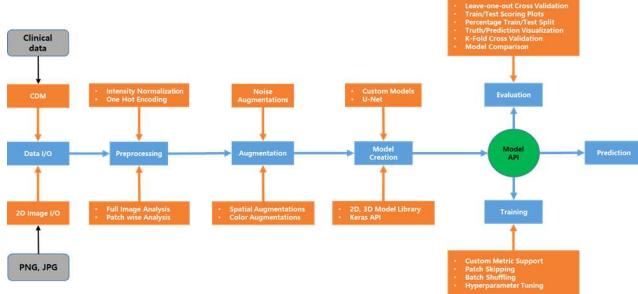
본 논문에서 제안한 플랫폼은 표준화된 의료빅데이터 인프라를 기반으로 하며, 의료영상은 포함한다. 그리고 이를 기반으로 관리 및 지원을 위한 대표적인 기능으로 보안 및 규칙기반 저장과 검색할 수 있는 인터페이스를 제공한다. 또한 의료데이터를 기반으로 개발 소프트웨어 도구와 분석 소프트웨어를 갖는다. (그림 1)은 의료빅데이터 기반 확장 인공지능 통합 플랫폼의 전체 구조를 보인다. 대표적인 질환 연구로는 근감소증과 간질환 그리고 요로

결석과 같은 복부영역에 중점을 두었다. 특히, 의료빅데이터 표준인 CDM과 의료영상 표준인 R-CDM을 통해 표준화된 데이터셋을 제공하도록 한다.



(그림 1) 제안한 플랫폼 구조

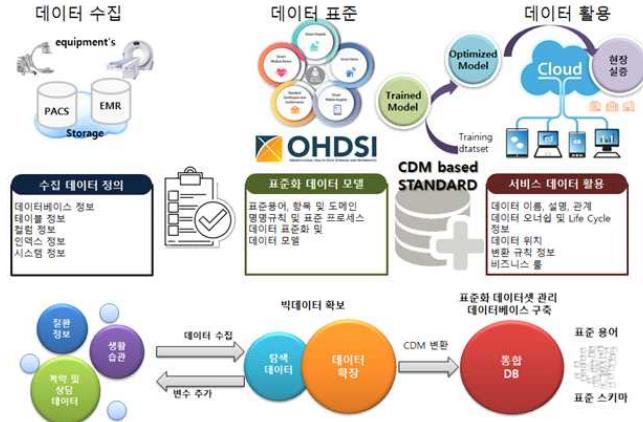
제안하는 플랫폼의 전체 기능은 (그림 2)와 같다. 전체 주요 기능으로 데이터 입출력, 전처리 그리고 데이터 증강 그리고 모델의 생성 그리고 이를 기반으로 평가와 학습 그리고 예측 모델을 개발하게 된다.



(그림 2) 머신러닝 개발 플랫폼의 전체 기능

3. 의료 빅데이터의 통합과 현장실증 활용

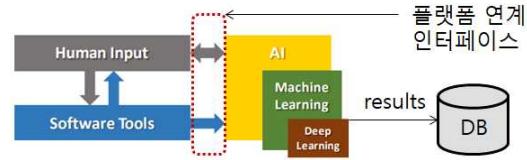
의료빅데이터는 병원정보시스템의 전자의무기록(EMR)과 PACS를 기반으로 한다. 이와 함께 질환정보, 생활습관 그리고 복약 및 상담데이터까지 포함하여 OMOP-CDM 기반의 공통데이터모델로 변환하여 저장한다. (그림 3)과 같이 학습용 데이터로 활용하며, 최적화 모델을 개발하여 현장 실증으로 활용하게 된다.



(그림 3) 의료빅데이터 통합과 활용

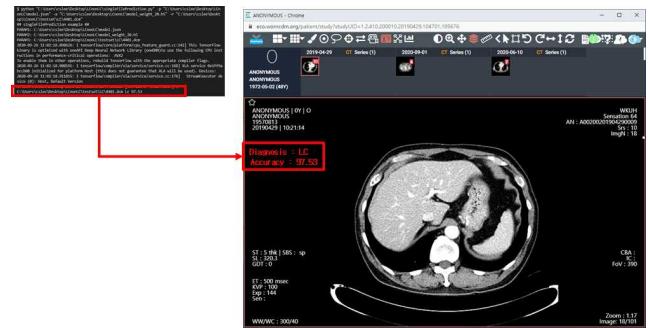
머신러닝 연구를 위한 처리과정은 데이터 수집과 표준화된 데이터를 기반으로 학습을 위한 데이터셋을 생성시

키는 기능이 가장 중요하며, 학습 데이터셋에 대한 어노테이션 기능을 제공해야 한다. 또한 어노테이션 기능은 매뉴얼 방식과 자동방식으로 제공한다. 사용자의 입력과 분석 소프트웨어 도구를 통해 레이블링 데이터를 생성하기 위한 인터페이스를 제공한다.



(그림 3) 머신러닝 기술개발 플랫폼의 역할

(그림 5)는 웹기반 의료데이터관리 플랫폼기반으로 인공지능 모듈을 적용시켜 간경변(cirrhosis of the liver) 질환 진단 결과를 보이고 있다.



(그림 5) 웹기반 인공지능 모듈 적용

4. 결론 및 향후 연구

본 논문에서는 의료빅데이터기반 화장 인공지능 통합 플랫폼을 제안한다. 이를위해 의료빅데이터 표준화 인프라를 기반으로 하며 표준화된 데이터셋을 기반으로 학습 데이터 모델링과 최적화 모델을 개발하여 현장에 활용되도록 지원할 수 있는 기능을 정리하였다. 현재 웹기반 플랫폼으로 각 기능별 개발이 진행되고 있고 향후 통합하여 원스탑 지원 플랫폼으로 활용할 계획이다.

향후 계획으로는 특정 질환 데이터의 수집과 표준화 작업 그리고 머신러닝 알고리즘을 개발하여 병원정보시스템과 연계하여 스크리닝 목적 또는 인공지능 기술개발의 타탕 성 검증 목적의 시스템 개발을 계획하고 있다.

사사표기

본 연구는 보건복지부의 재원으로 한국보건산업진흥원의 보건의료기술 연구개발사업(HI18C1216) 그리고 한국연구재단(NRF-2020R1I1A1A01074256)과 (NRF-2016M3A9A7918501) 지원에 의하여 이루어진 것임.

참고문헌

- [1] Haleem, A., Javaid, M., & Khan, I. H. (2019). Current status and applications of artificial intelligence (AI) in medical field: an overview. *Current Medicine Research and Practice*, 9(6), 231–237.

[2] MURDOCH, Travis B.; DETSKY, Allan S. The inevitable application of big data to health care. *Jama*, 2013, 309.13: 1351–1352.