대형 멀티모달 모델의 BIM 입찰문서 도표 분석 능력 조사

김정수*
*한국건설기술연구원 구조연구본부, BIM 클러스터
e-mail:jeongsookim@kict.re.kr

Investigation on Analysis Capacities of a Large Multimodal Model for Both Figures and Tables in BIM Bidding Documents

Jeongsoo Kim*
*Dept. of Structural Engineering, BIM Cluster, Korea Institute of Civil Engineering and Building Technology

요 약

건설사업의 입찰문서에 포함된 다양한 BIM 적용 요구사항을 분석하고 이해할 수 있는 인공지능 모델에 대한 수요가 높아지고 있다. 본 연구는 입찰문서 내 BIM 요구사항에 대한 양질의 데이터셋 구축을 위해 대형 멀티모달 모델의 BIM 입찰문서 도표에 대한 분석 능력을 조사하였다. BIM 건설사업에 대한 입찰문서로부터 구축한 도표 데이터셋에 사용 모델의 명시적, 내재적 정보 추출 능력을 분석할 수 있도록 프롬프트 명령어를 설계하고, 100여장의 도표에 대해 분석결과를 비교하였다.

1. 서론

2022년부터 1천억원 이상 공공건설 발주 사업에서 건설정보모델(Building Information Model, BIM) 사용이 점진적으로 의무화됨에 따라 건설 분야에서의 BIM 전문인력에 대한 수요가 높아지고 있다. 최근 발표된 입찰단계에서부터 건설 전 주기에 걸쳐 BIM 적용 계획을 요구한 가덕도 신공항 사업과 같이, 입찰단계에서부터 건설 생애주기를 고려한 BIM 적용 요구사항이 증가할 것으로 예상되므로, 입찰 시 입찰문서의 정보를 정확하고 효율적으로 분석함으로써 해당 건설사업에 대한 예상 업무 및 투입 자원을 판단하는 것이 건설 기업에게 중요한 역량이 될 것으로 전망된다. 그러나 일부 대형 건설회사를 제외하고 대부분의 건설회사는 관련 경험을 갖춘 인력 확보에 어려움을 겪고 있으며, BIM 사업의 요구사항에 대한 분석에 기반해 입찰사업에 참여할 수 있는 역량을 보유하지 못한 상황이다.

최근 발주-입찰단계에서 BIM 전문인력에 대한 건설 주체 간 불균형을 완화하기 위한 연구가 이뤄지고 있으며[1], 특히 입찰안내서, 제안요청서, 수행계획서 등의 BIM 입찰문서로부터 BIM 관련 텍스트를 인식하고 해당 텍스트의 BIM 요구사항에 대한 수행 난이도를 구분할 수 있도록 BIM 실무 전문가의 경험 지식을 학습한 딥러닝 모델이 제안되기도 하였다[2]. 그러나 대부분의 BIM 입찰문서에서 BIM과 관련된 많은 정보가 그림 및 표를 통해 제시되고 있어, 문서로부터 추출된 텍스트 정보만으로는 입찰문서를 기계적으로 분석하는데 한계가 있다. 그림 및 표에 대한 정보를 추출함으로써 BIM에 대한 다양한 요구사항을 파악할 수 있으며, 이를 학습시킨 딥러닝 모델을 활용함으로써 중소 건설회사의 BIM 경쟁력을 높일 수 있을 것으로 판단된다.

본 연구는 다양한 데이터 유형으로부터 데이터를 생성할 수 있는 대형 멀티모달 모델(Large Multimodal Model, LMM)의 BIM 입찰문서의 도표로부터 BIM 관련 정보를 추출 성능을 조사하였다.

2. 연구방법

본 연구는 공개 LMM 모델을 Lee et al. [3]에서 구축한 건설분야 입찰문서 이미지 데이터셋에 적용하고, 해당 모델의 추론 결과를 조사하였다.

도표 이미지 데이터셋으로부터 텍스트 정보를 추출하기 위해 공개 LMM 모델 중 하나인 Gemini-pro-vision (version 1.0) 모델을 사용하였다. 해당 모델은 Google DeepMind에서 개발된 생성형 인공지능 모델 중 하나로 멀티모달 프롬프트를 지원하고, 최신 LMM 모델은 아니나 별다른 과금없이

(2024년 4월 기준) 응용 프로그래밍 인터페이스를 제공하고 있어 다량의 이미지 분석에 활용할 수 있다.

100장의 도표 이미지에 대해 표 1의 프롬프트 명령어를 사용하여 생성된 결과를 각각 저장하였고, 이를 정성적으로 비교 분석하였다. 표 1의 프롬프트 명령어의 설계 의도는 그림에서의 텍스트 추출 성능(Case A), 표 자체에 대한 독해 능력(Case B), 구성 요소 간 관계 분석 등 잠재 정보 추출(Case C)을 염두하였다.

[표 1] 프롬프트 명령어셋

Case	명령어	사용모델
A	모든 텍스트를 추출하시오 + 다음 텍스트를 1개 문장으로 작성하시오:	Gemini-pro-vision + Gemini-pro
B	도표에서 직접적으로 제공하는 정보를 1개 문장으로 작성하시오	Gemini-pro-vision
C	도표로부터 추론할 수 있는 정보를 1개 문장으로 작성하시오	Gemini-pro-vision

3. 연구결과 및 결론

그림 1은 표 1의 명령어를 적용해 얻은 분석 결과 중 하나를 나타낸 것이다. 그림 1에서 확인할 수 있듯이, 사용된 LMM 모델은 이미지 내 포함된 모든 글자를 추출하는 등 우수한 글자 인식 성능을 보유하였으며(Case A), 도표에서 요구하는 과업수행조직 인원의 구성 조건도 제시해내었다(Case B). 또한 수행 조직도의 위계를 추정하였으나, 수행조직을 실제 총괄매니저 관리하에 코디와 모델러가 구성된 조직으로 분석하지 못하고, 매니저와 코디로 구성된 조직으로 분석하였다.

사용된 LMM은 여러 BIM 입찰문서의 도표에 대해 텍스트의 인식 및 요약 그리고 도표로부터 부가적으로 알 수 있는 정보를 추출할 수 있는 것으로 확인되나, 도표의 구성이 복잡해질수록 도표에 내재해 있는 추론 성능이 떨어졌으며, 도표 내 텍스트 정보가 많아질수록 텍스트가 자연스럽지 못한 경우가 다수 확인되었다. 최신 LMM 모델의 사용, 관련 데이터 전이 학습 등을 통한 최적화된 모델 (Tuned model)을 사용, 최적화된 프롬프트 설계 등을 수행함으로써 보다 양질의 BIM 요구사항 정보를 추출해 나갈 수 있을 것으로 판단된다.

본 연구는 LMM모델의 BIM 입찰문서 내 도표에 적용하여 얻은 분석결과로부터 BIM 관련 명시적, 잠재적 정보의 추출 가능성을 분석하였다. 추후 고성능의 LMM 모델의 최신 BIM 입찰문서에 적용해 BIM의 요구사항 및 기술 정보를 추출 성능을 다각도로 분석하고자 하며, 향후 양질의 BIM 입찰 요구사항 데이터셋을 구축해 건설분야 BIM 입찰 관련 딥러닝 모델 개발 등 후속 연구를 수행하고자 한다.

감사의 글

이 논문은 국토교통부 및 국토교통과학기술진흥원의 국가 R&D “BIM 기반 인프라 발주-설계 프로세스 디지털 협업 체계 개발사업” (RS-2023-00143371)의 연구비 지원을 받았습니다. 연구비 지원에 감사드립니다.

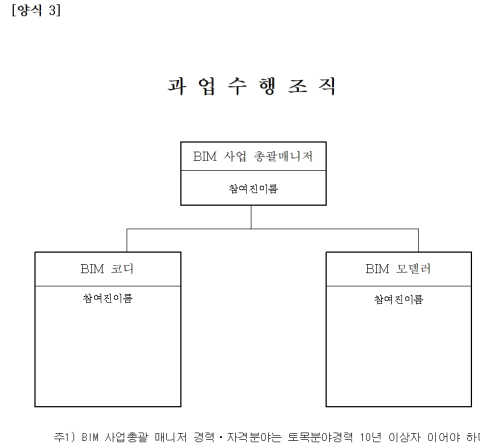
참고문헌

[1] 문현석, “BIM기반 디지털 발주-입찰-평가 프로세스 통합 모형 개발”, 대한토목학회 정기학술대회, 부산, pp.1-2, 10월, 2022년

[2] 김정수, 문현석, 박상미, “건설사업 입찰 텍스트의 BIM 수행 난이도 추론을 위한 딥러닝 모델”, 대한토목학회논문집, 제 43 권 6호, pp.851-863, 12월, 2023년

[3] 이혜민, 김정수, 문현석, “딥러닝 객체 인식을 활용한 문서 내 표 텍스트 정보 추출”, 산학기술학회 춘계학술대회, pp.531-532, 5월, 2023년

[그림 1] 프롬프트 명령어 적용결과 일례



Case	LMM 분석 결과
A	[양식 3] 과업수행 조직 BIM 사업 총괄매니저 참여진이름 BIM 코디 참여진이름 BIM 모델러 참여진이름 주1) BIM 사업총괄 매니저 경력 : 자격분야는 토목분야이며 경력 10년 이상자이며 하며, BIM 코디는 BIM 분야 경력 5년 이상자이며 한다. 또한 BIM 모델러는 BIM기술 자로 2인 이상 구성하여 한다
	BIM 사업 총괄매니저(토목분야 10년 이상 경력), BIM 코디(BIM 분야 5년 이상 경력), BIM 모델러(BIM 기술자 2인 이상)로 구성된 과업수행 조직.
B	도표에서 직접적으로 제공하는 정보는 "BIM 사업은 매니저 경력 · 자격분야는 토목분야이며 경력 10년 이상자이며 하며, BIM 코디는 BIM 분야 경력 5년 이상자이며 이어주가 있으며, 또한 BIM 모델러는 BIM기술 자격증을 취득한 자로 2인 이상 구성하여 한다."입니다.
C	[양식 3] 과업수행 조직도에서 작업자는 팀장과 모델러로 구성되어 있으며, 팀장은 코디와 매니저로 구성되어 있다.