

생성형 AI의 건설산업 활용

연세대학교 건설환경공학과
스마트인프라연구실
조교수 김홍조

2024.10.25



목차

1. Foundation Models in CV
2. RAG-GPT



Foundation Models in Computer Vision

ChatGPT, Gemini

OpenAI Menu

ChatGPT

ChatGPT

Get instant answers, find creative inspiration, learn something new.

Download on the App Store | GET IT ON Google Play

[Try on web](#)

Overview | Team Enterprise | ChatGPT 3.5

Me
I am hosting my friends for brunch on Sunday and they are asking what to bring. Could you give me a list of suggestions? I'll be making a frittata and mimosas.

ChatGPT
Certainly! Here's a list of brunch items your friends could bring:

1. Fresh fruit salad or a fruit platter
2. Pastries or croissants
3. Bagels with cream cheese and smoked salmon
4. Yogurt with granola and berries
5. Breakfast potatoes or hash browns
6. Sausages or bacon
7. Quiches or savory tarts
8. Pancakes or waffles with syrup
9. Coffee or tea
10. Freshly squeezed orange juice

These options can complement your frittata and mimosas nicely and provide a variety of flavors for your brunch gathering.

Gemini

+ New chat

Hello, Hongjo.
How can I help you today?

ⓘ Gemini Apps Activity is off
Some things that require Gemini Apps Activity are also turned off, like pinned & recent chats and extensions. Turning off Gemini Apps Activity doesn't delete past data.

- As a social trend expert, explain a term
- Recommend new types of water sports, including pros & cons
- Suggest the best parks to visit in a city with descriptions
- Brainstorm team bonding activities for our work retreat

Help | Activity | Settings | Upgrade to Gemini Advanced

South Korea, Gyeonggi-do, Goyang-si, Deogyang-gu, 행주동

Enter a prompt here

ChatGPT?

The screenshot shows the ChatGPT web interface. On the left, there is a sidebar with navigation options: 'ChatGPT', 'DALL-E', and 'Explore GPTs'. The main area displays a dropdown menu for 'ChatGPT 3.5' with three options: 'GPT-4' (with DALL-E, browsing and analysis, limit 40 messages / 3 hours), 'GPT-3.5' (Great for everyday tasks, selected), and 'Plugins'. Below the menu is the OpenAI logo and the text 'How can I help you today?'. There are four suggested prompts: 'Recommend a dish to bring to a potluck', 'Plan an itinerary for a fashion-focused exploration of Paris', 'Make up a story about Sharky, a tooth-brushing shark superhero', and 'Give me ideas for what to do with my kids' art'. At the bottom, there is a text input field 'Message ChatGPT...' with an upward arrow, a settings gear icon, a share icon, and a character count '99999 chars'. A disclaimer at the bottom reads 'ChatGPT can make mistakes. Consider checking important information.'

- ChatGPT는 OpenAI에 의해 개발된 인공지능 기반의 대화형 모델
- 대량의 텍스트 데이터를 학습
- 사람과 유사한 방식으로 텍스트를 생성하고 대화할 수 있도록 설계
- 사용자의 질문이나 명령에 대해 자연스러운 언어로 응답하며, 다양한 주제에 대한 정보를 제공하고, 특정 작업을 수행하는 데 도움

GPT-4 (2023.3)

- 멀티 모달 모델(이미지 및 텍스트 허용)
- 모의 변호사 시험에 합격하여 응시자 중 상위 10%의 점수 획득
- 동일한 변호사 시험에서 GPT-3.5는 하위 10%의 점수 획득

GPT-4 Technical Report

OpenAI*

Abstract

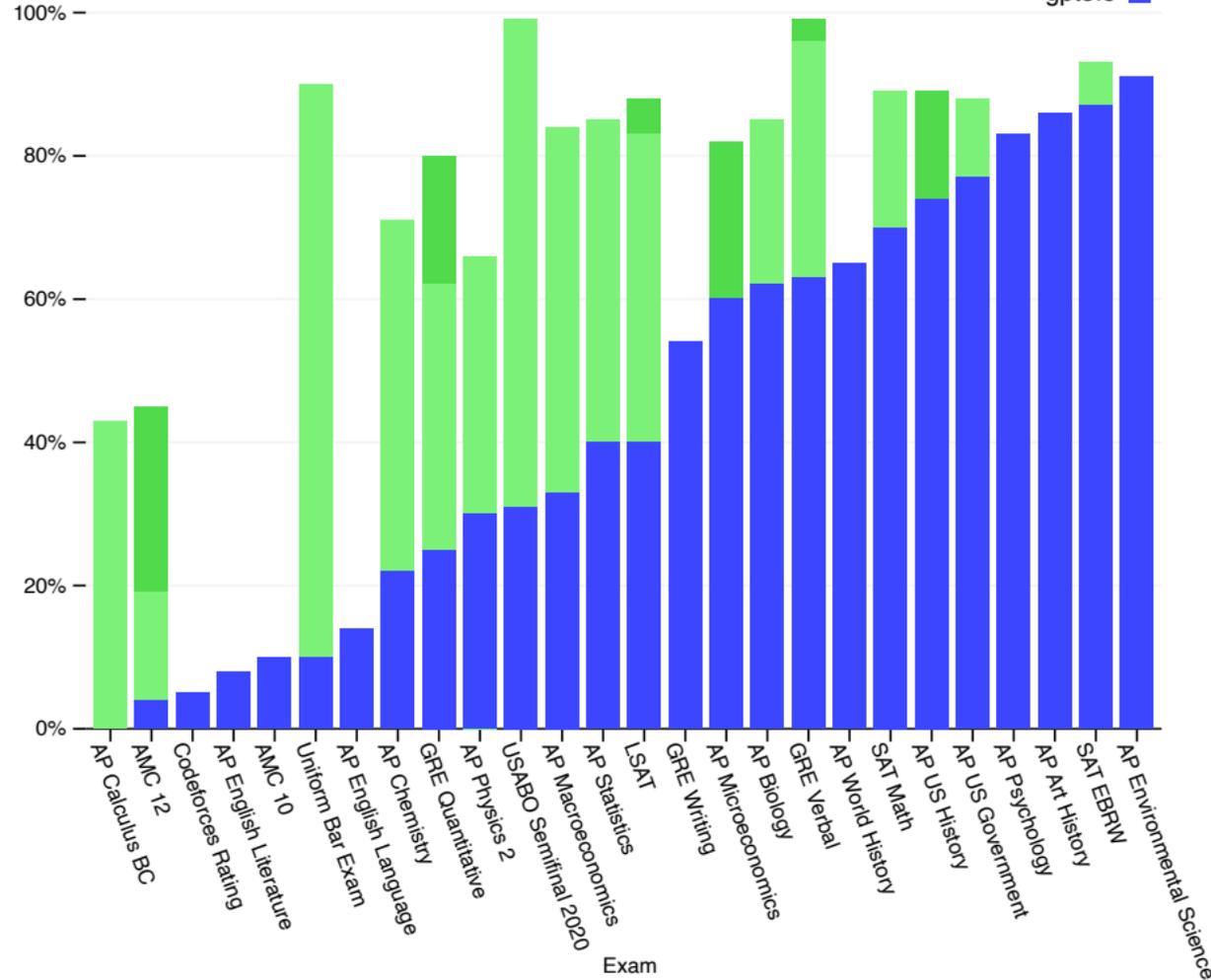
We report the development of GPT-4, a large-scale, multimodal model which can accept image and text inputs and produce text outputs. While less capable than humans in many real-world scenarios, GPT-4 exhibits human-level performance on various professional and academic benchmarks, including passing a simulated bar exam with a score around the top 10% of test takers. GPT-4 is a Transformer-based model pre-trained to predict the next token in a document. The post-training alignment process results in improved performance on measures of factuality and adherence to desired behavior. A core component of this project was developing infrastructure and optimization methods that behave predictably across a wide range of scales. This allowed us to accurately predict some aspects of GPT-4's performance based on models trained with no more than 1/1,000th the compute of GPT-4.

GPT-4 (2023.3)'s performance in academic and professional exams

Exam results (ordered by GPT-3.5 performance)

Estimated percentile lower bound (among test takers)

gpt-4
gpt-4 (no vision)
gpt3.5



GPT-4 (2023.3)'s reasoning capability

GPT-4 visual input example, Extreme Ironing:

User What is unusual about this image?



Source: <https://www.barnorama.com/wp-content/uploads/2016/12/03-Confusing-Pictures.jpg>

GPT-4 The unusual thing about this image is that a man is ironing clothes on an ironing board attached to the roof of a moving taxi.

GPT-4 (2023.3)'s reasoning capability

GPT-4 visual input example, Chicken Nugget Map:

User Can you explain this meme?

Sometimes I just look at pictures of the earth from space and I marvel at how beautiful it all is.



GPT-4

This meme is a joke that combines two unrelated things: pictures of the earth from space and chicken nuggets.

The text of the meme suggests that the image below is a beautiful picture of the earth from space. However, the image is actually of chicken nuggets arranged to vaguely resemble a map of the world.

The humor in this meme comes from the unexpected juxtaposition of the text and the image. The text sets up an expectation of a majestic image of the earth, but the image is actually something mundane and silly.

of Parameters

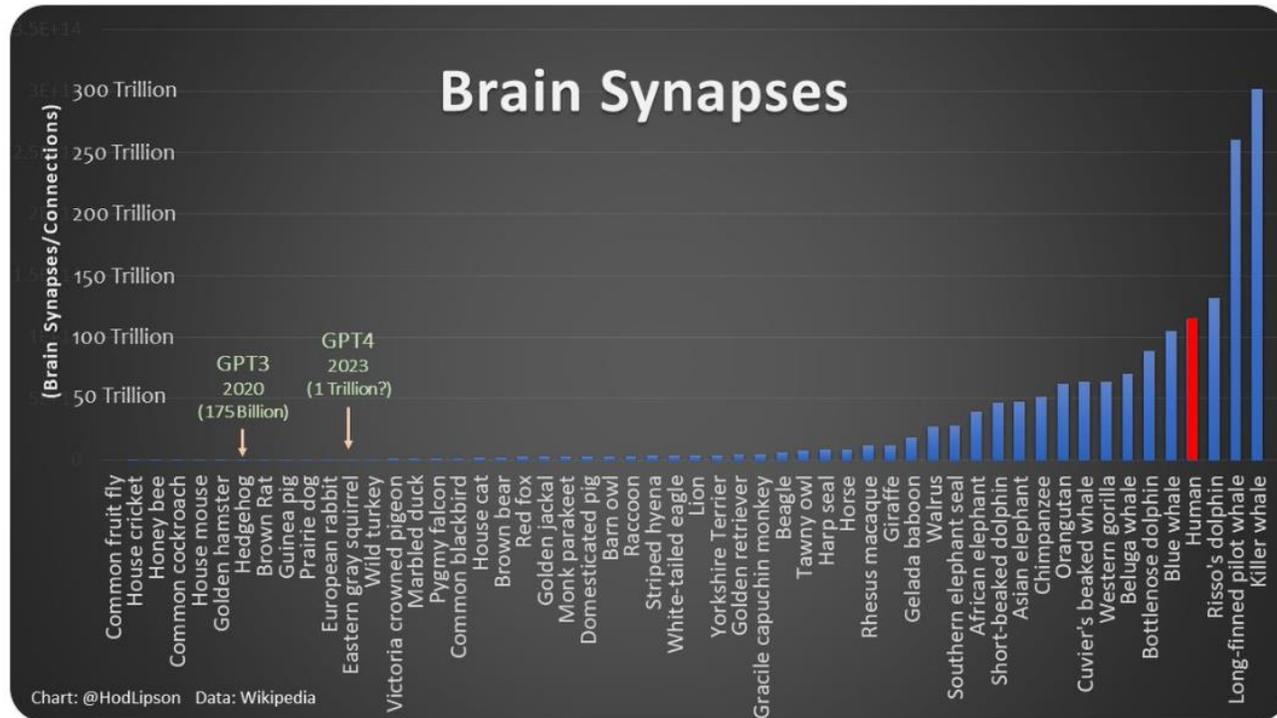


Hod Lipson 
@hodlipson

팔로우 ...

@OpenAI's #GPT3 has 175 billion synapses (like a hedgehog). Next year's #GPT4 is rumored to have about 1 trillion (like a squirrel). By 2030, #GPT7 may have as many synapses as a human brain (~120 trillion) -- if only we can figure out how to train it.

[게시물 번역하기](#)

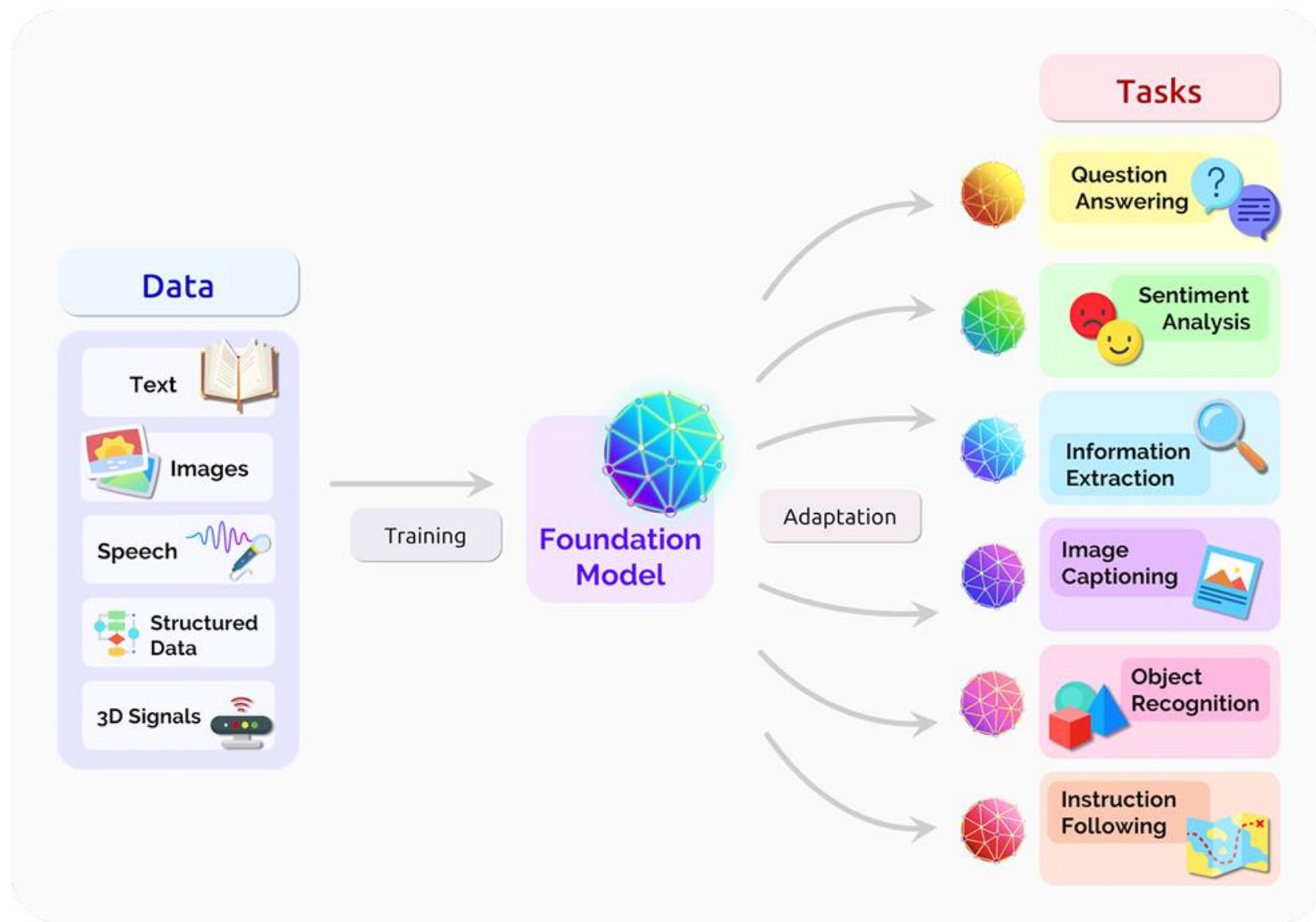


건설 산업에서 생성형 AI의 활용 전망

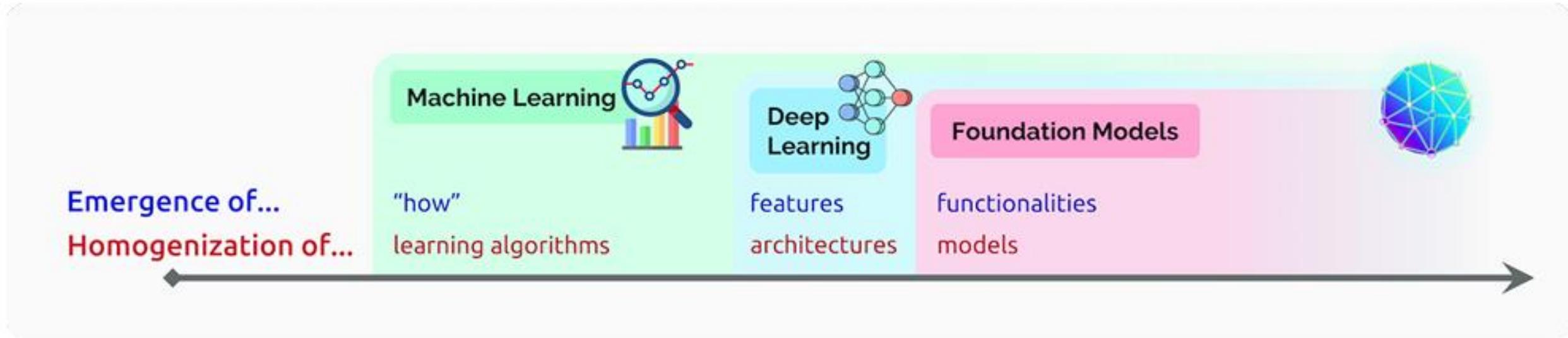
- 설계 최적화: AI를 통한 설계 최적화
- 프로젝트 관리: 비용과 일정 예측, 리스크 관리 개선
- 고객 서비스 자동화: 자동 응답 및 맞춤형 건축 솔루션 제안
- 안전 및 위험 관리: 사고 예방을 위한 위험 식별 및 대응 계획 제시
- 지식 공유 및 교육: 경험과 **Best practic**의 대화형 공유 및 학습 촉진
- 전문사무의 자동화: 엔지니어링 문서 생성 업무 자동화를 통한 생산성 향상
- 문제 발견 및 해결책 제시 (**GPT-5 이후**): 큰 문제를 작은 문제로 나누어 각각의 해결책 또는 대안을 제시

Foundation Model (범용 모델)

- 이 모델들은 다양한 작업에 적용할 수 있는 범용적인 AI 모델
- 이들은 대규모의 데이터셋에서 학습되며, 다양한 작업에 적용될 가능



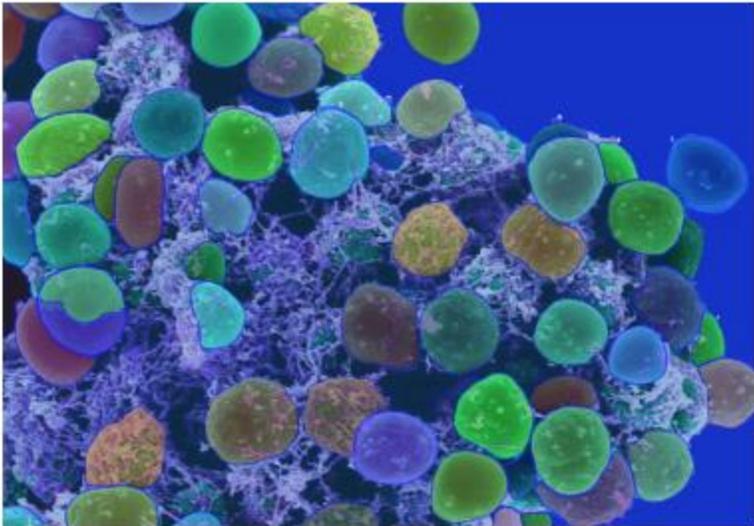
AI 기술의 출현 및 동질화 (emergence and homogenization)



- 머신러닝이!도입되면서!예제에서!작업!수행!방식이!등장)자동!추론*
- 딥러닝을!통해!예측에!사용되는!고급!기능이!등장
- 기반!모델을!통해!상황에!맞는!학습과!같은!고급!기능까지!등장
- 머신러닝은!학습!알고리즘을!동질화)예;!경사!하강법*!
- 딥러닝은!모델!아키텍처를!동질화)예;!컨볼루션!신경망*
- 기초!모델은!모델!자체를!동질화)예;!HQU. 4*

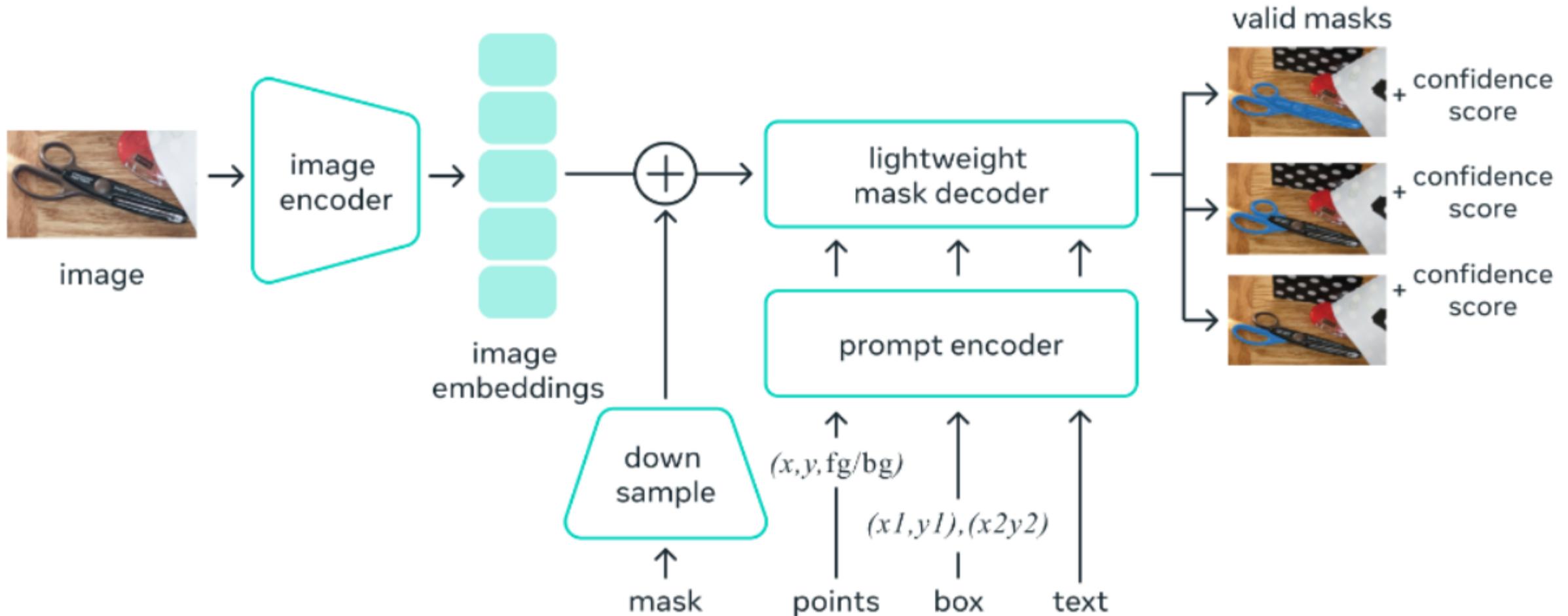
범용 객체 분할 모델 SAM (Segment Anything Model)

- Segment Anything 모델은 이미지 내 객체를 분리하는 데 사용
- 11백만 개의 이미지와 11억 개의 마스크에 대해 학습
- 사용자가 제공하는 프롬프트를 사용하여 모든 객체 마스크 생성 가능



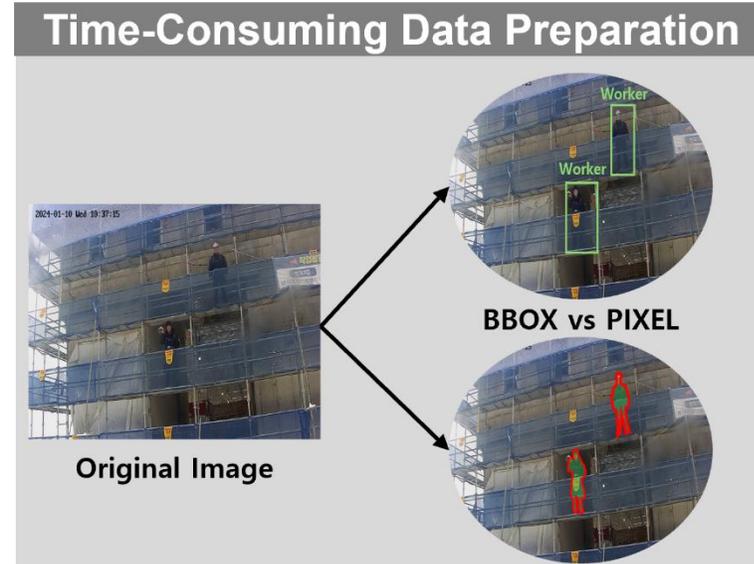
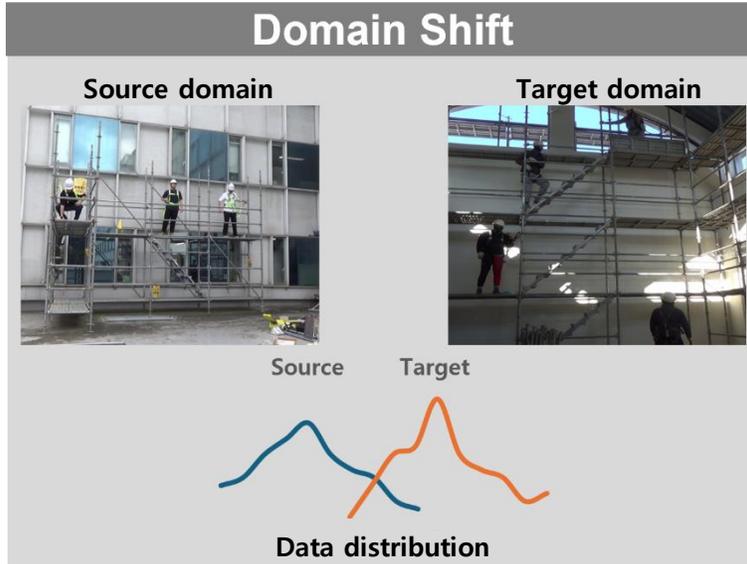
범용 객체 분할 모델 SAM (Segment Anything Model)

Universal segmentation model



Problem statement & Goal

Computer Vision Task



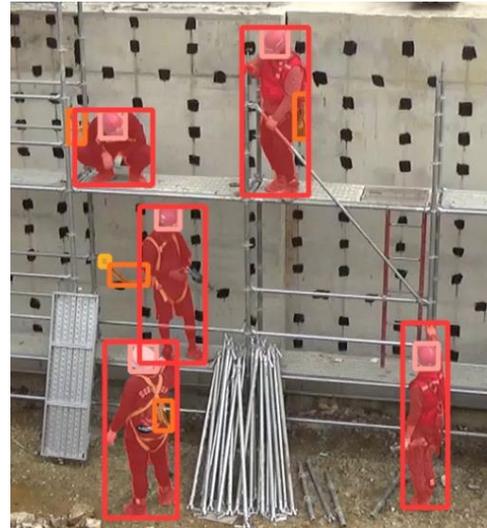
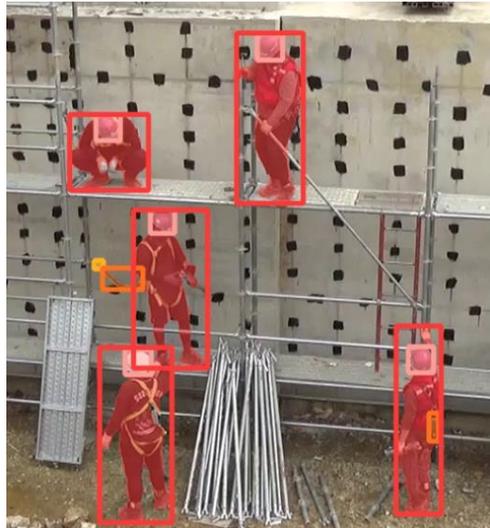
- **Distribution difference**(color, background..) between the source domain and target domain
- **Time-consuming** process of preparation labeling data
- **Detection challenges for small objects** like personal protective equipment (strap, hook..)

Causes of Performance Degradation in Deep Learning Models

Improving the Performance of Deep Learning Models

Segmentation model performance

Target1	Mask mAP@50	
	Baseline model	Optimized model
Worker	49.9	86.2
Hardhat	61	99.2
Strap	0	46.8
Hook	0	56.6



Target2	Mask mAP@50	
	Baseline model	Optimized model
Worker	99.1	97.3
Hardhat	61	94.3
Strap	59.7	75.8
Hook	57.7	71.8

Findings

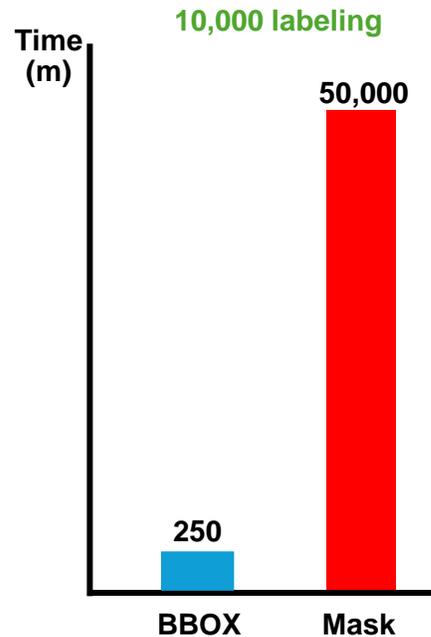
Reduce mask label data generation time



➔ **30 to 90**
seconds
per image



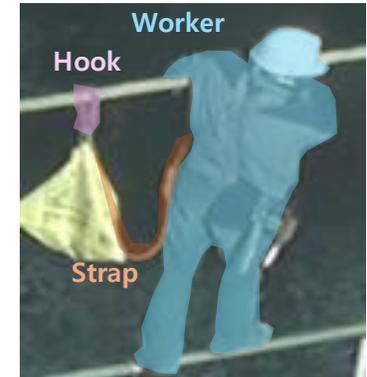
➔ **3 to 5**
minutes
per image



Improving object performance



Before



After

Improving small object detection performance

If there are more images and classes?

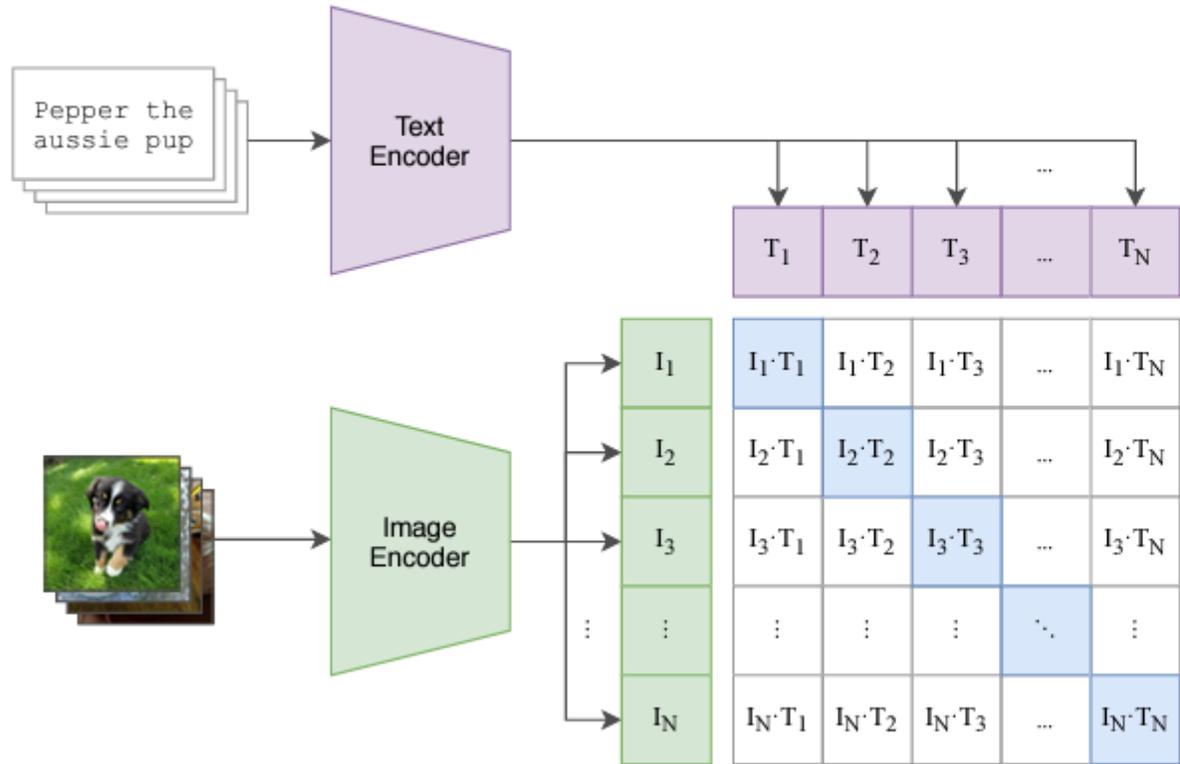
If the BBOX labels are manually generated by a person, then the mask labels can be automatically generated!

CLIP(Contrastive Language-Image Pre-Training) 모델

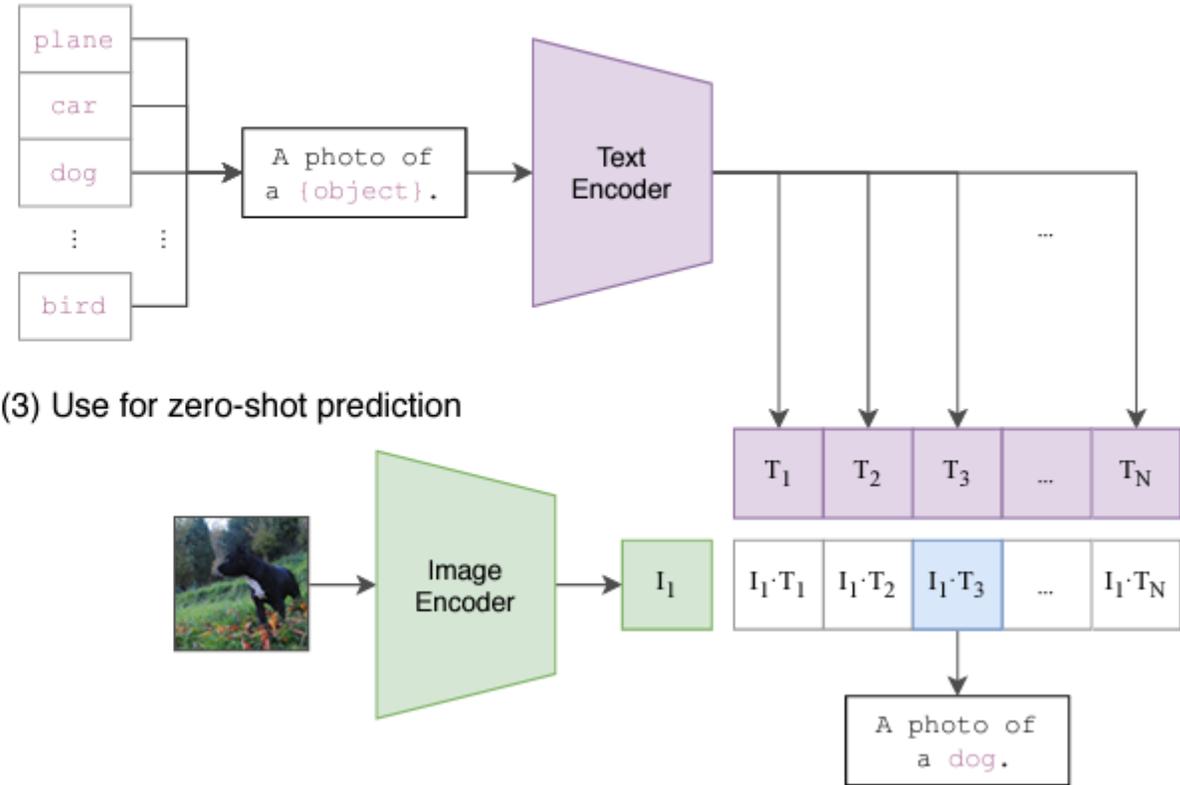
- CLIP 모델은 이미지와 텍스트 캡션 사이의 관계를 학습하는 데 사용
- 이 모델은 특정 작업에 대해 직접 최적화하지 않고도 이미지에 가장 관련성 있는 텍스트 예측
- 이 모델은 원래 ImageNet에서 128만개의 레이블이 붙은 예제를 사용하지 않고도 ResNet-50에 비견, 그 이상 성능을 "제로샷"으로 확보

CLIP(Contrastive Language-Image Pre-Training) 모델

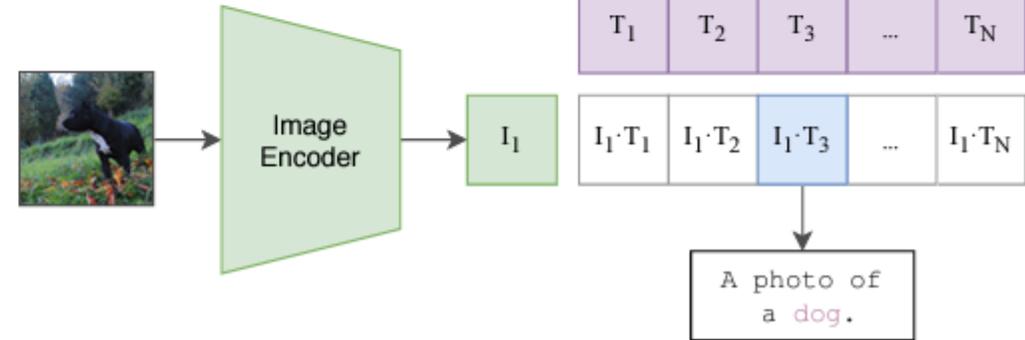
(1) Contrastive pre-training



(2) Create dataset classifier from label text

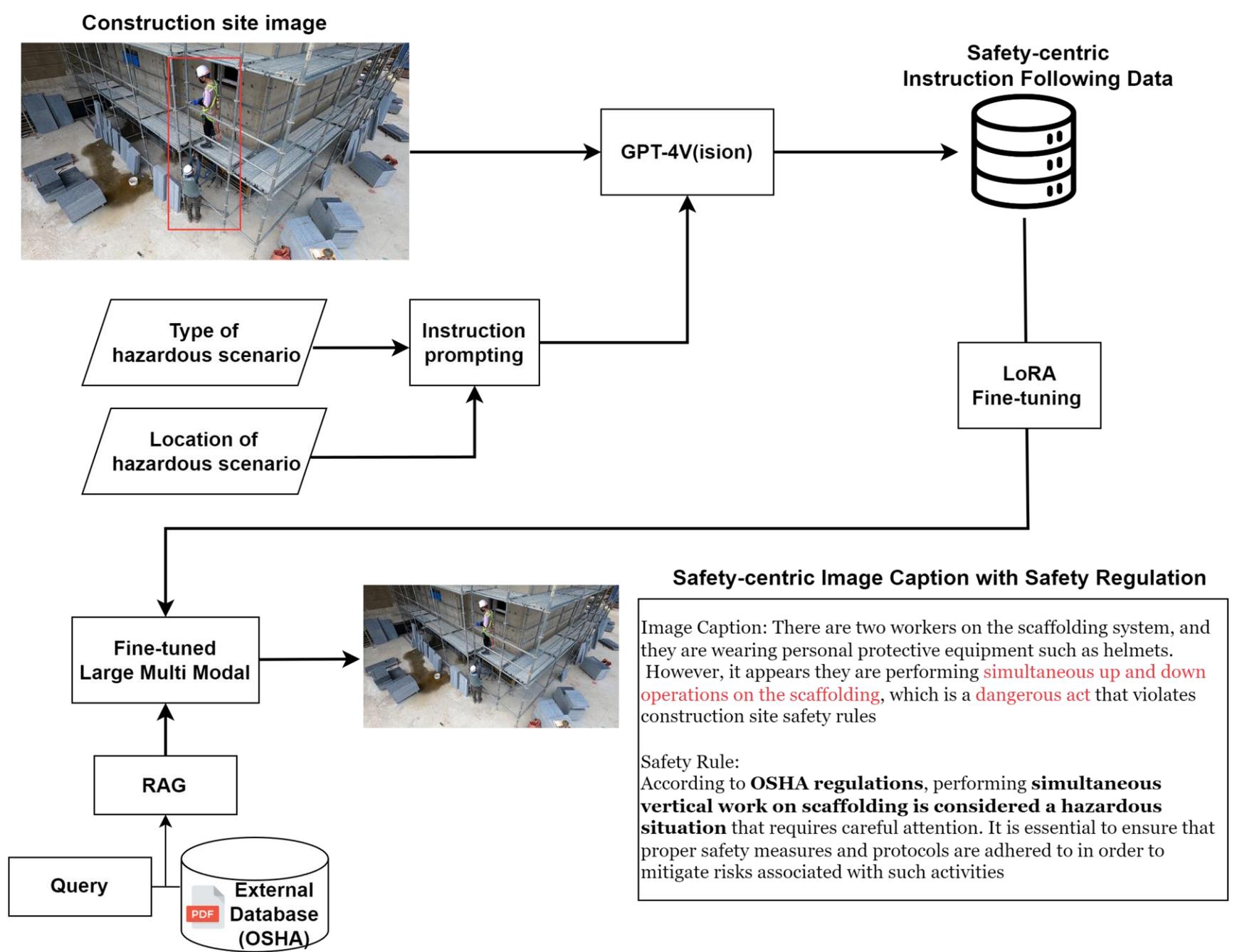


(3) Use for zero-shot prediction



- 표준!이미지!모델은!이미지!특징!추출기와!선형!분류기를!함께!훈련하여!라벨을!예측하는!반면!DMQ은!이미지!인코더와!텍스트!인코더를!함께!훈련하여!이미지!텍스트*!훈련!예제의!올바른!쌍을!예측

Large Multimodal Model for Context-aware Construction Safety Monitoring



Large Multimodal Model for Context-aware Construction Safety Monitoring

Multi-Task

Examples of Context-aware Construction safety Monitoring



Look at the image in detail
Answer for following 2 requirements

1. Provide the description of the image
2. Please respond with either 'Safe' or 'Unsafe' to indicate whether the situation in the photo is dangerous or not

Proposed fine-tuned model :

1. A forklift is depicted in the image, carrying a stack of materials that **appears to obstruct the driver's forward visibility**, posing a potential safety hazard. A worker is standing nearby, wearing a safety vest and hard hat, indicating a commitment to safety protocols despite the unsafe material loading practice.
2. **Unsafe**. The materials loaded on the forklift **obstruct the driver's line of sight**, creating a risk of accidents due to reduced visibility.

GPT-4V:

1. The image shows an outdoor setting where a person in a high-visibility vest and hardhat is standing on a paved area, observing a forklift. The forklift is transporting a large stack of white material which appears to be securely placed on its forks. Behind the scene is a fence and beyond that, some dense greenery. The worker is standing a safe distance away from the forklift, which is currently stationary.
2. **Safe**, The situation seems to be under control, with safety measures being followed as the worker maintains a safe distance from the operational forklift. The area around the forklift is clear, without evident hazards, and the load being carried by the forklift appears to be stable and properly secured.

Detailed-description (Image Captioning)

- Detailed description including construction object
- Context-aware in construction site

Complex Reasoning

- Complex Reasoning for safety assessment
- Context-aware in construction site

Large Multimodal Model for Context-aware Construction Safety Monitoring

Few shot: Detailed description



이미지 속에는 굴삭기가 중앙에 배치되어 있고, 작업 중인 것처럼 버킷을 들어 올리고 약간 기울어진 공사 현장을 관찰할 수 있습니다. 근처에 안전모와 눈에 잘 띄는 조끼를 착용한 복장으로 신호수임을 알 수 있는 사람이 한 명 있습니다. 신호수는 굴삭기와 안전한 거리 내에 적절히 배치되어 있으며, 옆이 기계를 향하도록 서 있어 작업을 지시하는 데 적극적으로 참여하고 있음을 알 수 있습니다. 굴삭기는 배경에 나뭇잎으로 둘러싸인 공터에 위치하고 있으며, 현장 가장자리를 따라 다양한 건설 자재가 방수포로 덮여 있습니다. 굴삭기 작업 중 신호수의 배치와 관련하여 모든 안전 프로토콜이 준수되고 있는 것으로 보입니다. 굴삭기 운전자와 신호수 사이의 명확한 통신 경로는 진행 중인 건설 활동을 위한 안전한 환경을 조성합니다.



현장 설명

- 중앙에 위치한 굴삭기
 1. 굴삭기는 중앙에 배치되어 있으며, 작업 중으로 버킷을 들어 올리고 약간 기울어져 있습니다.
- 신호수
 1. 안전모와 눈에 잘 띄는 조끼를 착용하여 신호수임을 알 수 있습니다.
 2. 신호수는 굴삭기와 안전한 거리를 유지하며 적절히 배치되어 있습니다.
 3. 신호수는 옆이 굴삭기를 향하도록 서 있어 작업 지시에 적극적으로 참여하고 있습니다.
- 배경 및 환경
 1. 굴삭기는 나뭇잎으로 둘러싸인 공터에 위치하고 있습니다.
 2. 현장 가장자리를 따라 다양한 건설 자재가 방수포로 덮여 있습니다.
- 안전 프로토콜
 1. 굴삭기 작업 중 신호수의 배치와 관련하여 모든 안전 프로토콜이 준수되고 있습니다.
 2. 굴삭기 운전자와 신호수 사이의 명확한 통신 경로가 유지되어 건설 활동의 안전을 보장합니다.

Large Multimodal Model for Context-aware Construction Safety Monitoring

Few shot: Detailed description



이 이미지에서는 두 명의 작업자가 비계 시스템에서 동시에 작업하는 건설 현장을 관찰하고 있는데, 이는 안전에 문제가 될 수 있습니다. 위쪽에 있는 작업자는 구조물의 약 1/3 지점에 있는 비계 위에서 있습니다. 흰색 헬멧과 눈에 잘 띄는 조끼를 착용하고 있으며, 땅을 내려다보며 비계를 잡고 있는 것으로 보입니다. 그 아래, 대략 비계 아래쪽에는 두 번째 작업자가 더 낮은 자세로 작업하고 있습니다. 이 작업자는 쪼그리고 앉아 있으며 흰색 헬멧과 안전대 등 안전 장비를 착용하고 있으며 지상에서 작업에 집중하고 있는 것으로 보입니다.

위와 아래의 서로 다른 수직 구역에서 두 작업자가 동시에 작업하는 것은 낙하물의 잠재적 위험이나 비계의 부주의한 움직임으로 인해 두 사람의 안전에 영향을 미칠 수 있기 때문에 위험할 수 있습니다. 특히 동일한 비계 구조물에서 서로 다른 높이에서 작업이 진행되는 경우 사고를 예방하기 위해 안전 프로토콜은 일반적으로 이러한 시나리오에서 엄격한 조정과 의사소통을 요구합니다. 명백하게 조율된 안전 조치 없이 작업이 밀접하게 이루어지는 것은 두 작업자의 안전을 보장하기 위해 즉시 해결해야 하는 안전하지 않은 작업 시나리오를 나타냅니다.

현장 설명

- 작업자들의 위치
 1. 두 명의 작업자가 비계 시스템에서 동시에 작업 중.
 2. 위쪽 작업자: 구조물의 약 1/3 지점에 있는 비계 위에서 있음.
 3. 아래쪽 작업자: 비계 아래쪽에 쪼그리고 앉아 있음.
- 작업자의 안전 장비
 1. 위쪽 작업자: 흰색 헬멧과 눈에 잘 띄는 조끼 착용. 비계를 잡고 땅을 내려다 보고 있음.
 2. 아래쪽 작업자: 흰색 헬멧과 안전대 등 안전 장비 착용. 지상에서 작업에 집중하고 있음.
- 안전 문제
 1. 두 작업자가 서로 다른 수직 구역에서 동시에 작업하는 것은 안전에 문제가 될 수 있음.
 2. 낙하물의 잠재적 위험: 위쪽 작업자의 작업으로 인해 아래쪽 작업자에게 위험이 생길 수 있음.
 3. 비계의 부주의한 움직임: 비계의 움직임이 두 작업자의 안전에 영향을 미칠 수 있음.
- 안전 프로토콜
 1. 동일한 비계 구조물에서 서로 다른 높이에서 작업하는 경우, 안전 프로토콜은 엄격한 조정과 의사소통을 요구.
 2. 명백하게 조율된 안전 조치 없이 작업이 밀접하게 이루어지는 것은 두 작업자의 안전을 보장하기 위해 즉시 해결해야 하는 문제.
- 결론
 1. 현재 작업 방식은 낙하물 위험 및 비계의 부주의한 움직임으로 인해 두 작업자의 안전에 심각한 위험을 초래할 수 있음.
 2. 안전한 작업 환경을 위해 조정된 안전 조치와 명확한 의사소통이 필요함.

Large Multimodal Model for Context-aware Construction Safety Monitoring

GPT-4V's Output example of Complex reasoning task

질문: 작업자가 머리 위 작업을 완료하기 위해 사다리의 맨 꼭대기에 서 있는 상황에서 이 자세와 관련된 즉각적인 위험은 무엇이며, 이러한 불안정한 관행을 시정하고 작업자의 안전을 보장하기 위해 어떤 조치를 취해야 하나요?

답변:

- 즉각적인 위험
 1. 추락 위험: 작업자가 균형을 잃고 추락할 가능성이 큼.
 2. 부상 가능성: 심각한 부상으로 이어질 수 있음.
 3. 설계 문제: 맨 위 사다리는 사람의 체중을 지탱하거나 작업 플랫폼으로 사용하도록 설계되지 않음.
 4. 높은 무게 중심: 작업자의 무게 중심이 높아져 추락 위험이 증가.
 5. 구조적 지지대 부족: 추락을 방지할 수 있는 구조적 지지대가 없음.
- 안전 조치
 1. 안전한 위치로 이동: 작업자는 즉시 사다리에서 더 낮고 안정적인 위치로 내려와야 함.
 2. 3점 접촉 유지: 작업자는 항상 사다리와 3점 접촉(한 손과 두 발, 또는 두 손과 한 발)을 유지해야 함.
 3. 사다리 안정화: 다른 작업자가 사다리 바닥을 잡아 사다리가 미끄러지거나 움직이지 않도록 해야 함.
 4. 적절한 작업대 사용: 높은 곳에서 작업할 때는 가드레일이나 안전하고 안정적인 접근 수단이 있는 높은 작업대나 비계를 사용하는 것이 이상적임.
 5. 사다리 사용 지침 준수: 작업에 적합한 유형의 사다리를 사용하고, 제조업체의 지침을 따르는 것이 중요함.



결론

- 범용 모델은 **Zero-Shot, Few-Shot** 학습 기반 역량이 탁월
- 범용 모델 기반 컴퓨터 비전 기술들이 기존의 딥러닝 모델들을 대체할 것
- 범용 모델이 최적화된 **CNN** 또는 **Transformer** 모델 수준의 성능을 보이기 위해서는 추가적 학습 데이터 구축 필요
- 범용 모델 전용 학습 데이터는 기존의 **Bounding Box** 또는 **Polygon Mask** 형식의 학습 데이터와 다른 이미지 캡션 형태
- 학습 데이터 구축에 필요한 노력이 상당부분 경감되어 관련 연구 진행 및 상용 서비스 개발 속도 가속

거대 언어모델 기반
대화형 건설안전 정보 제공 서비
스 구축

연구 배경

- **작업 유형에 따른 맞춤형 건설안전 교육 자료의 생성 필요 증가**
 - : 산업안전보건법이 전면 개정, 도급인과 사업주 모두에게 근로자에 대한 안전 및 보건 조치 의무 강화
- **작업 유형에 따른 맞춤형 건설안전 교육 자료 생성에 소요되는 시간과 비용 소요의 어려움**
 - : 산업안전보건법, 중대재해처벌법, 건설산업기본법, 건설기술진흥법 등 참고 자료 분산
- **거대 언어 기반 AI 챗봇 기술 도입의 잠재성 주목**
 - : 기존의 건설안전 자료 생성 시 소요되는 시간과 비용의 문제를 해결 가능성, 동시에 편향된 지식 혹은 허구의 정보 생성의 한계 (Javaid Mohd et al., 2023).
- **초거대 언어모델에 Reference 제공, Reference 기반 대화형 건설안전 정보 제공 서비스 구축**
 - : 정보 제공의 정확성 개선이 요구 (Oviedo-Trespalacios Oscar, et al., 2023)에 부응.

연구 범위 및 방법

연구범위

- 상해의 정도가 가장 높은 추락 사고로 한정
- 추락 사고 가능성이 높은 외부 벽면 비계설치 작업으로 범위 설정

연구방법

1) LLMs References Embedding

- NCS - 가설 공사 중 비계 작업 관련 필수 지식, 기술, 태도(KSA) 정보
- KOSHA - 가설 공사 작업 안전 지침
- 건설기술진흥법 시행령; 가설구조물의 구조적 안전성, 소규모 안전관리계획의 수립 기준
- KOSHA Case - 추락사고사례

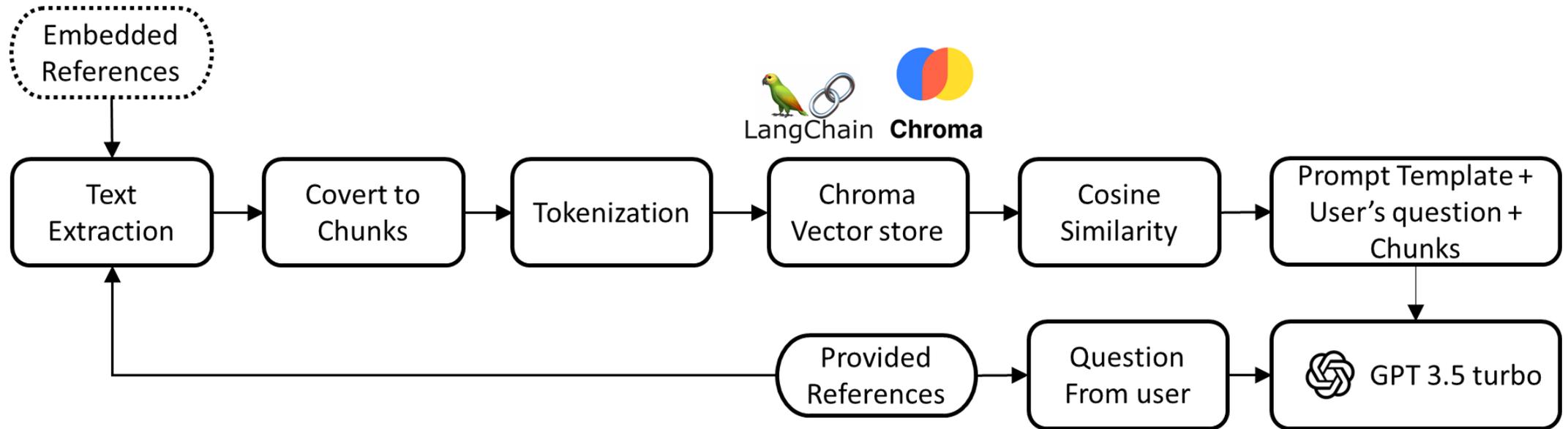
2) User-Specific Data + Question Import

- J건설 공사일보 전체 공정 대상 적용

3) Response

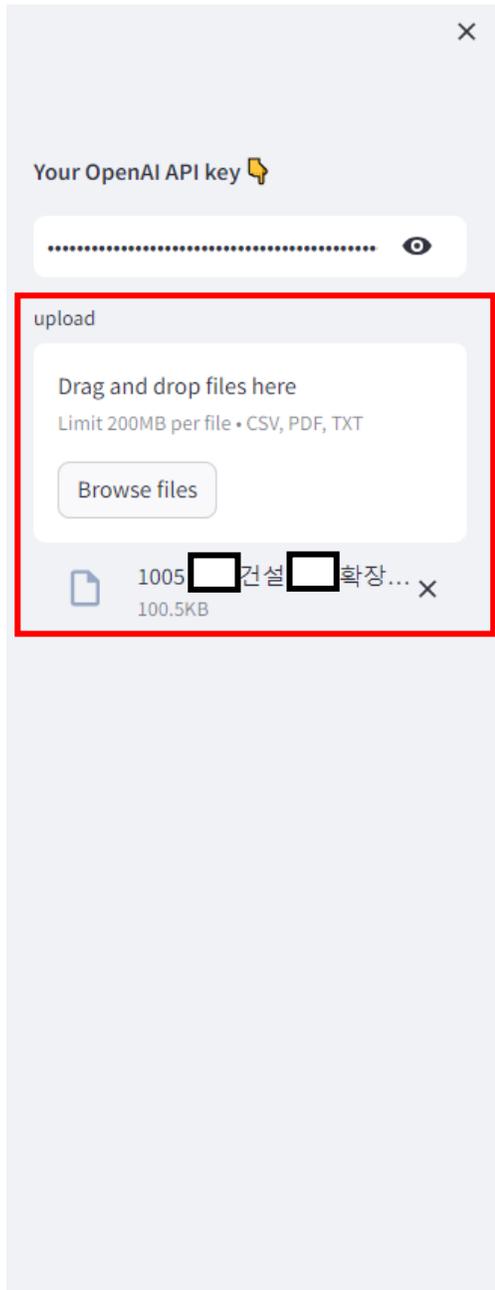
- 작업일보 기반 안전 작업 지침, 교육 내용 및 유사사례 정보 제공
- 건설안전관리자, 작업자 모두 안전 작업내용에 대한 질의 및 응답 가능

2.2 대화형 건설안전 정보제공 서비스 프로세스 제안



- Embedded References: 법령정보, 교육정보, 작업지침, 및 최근 사고 사례
- Text extraction: reference로부터 문단 취득
- Convert to chunks: 취득 문단을 문장 단위로 재구성
- Tokenization: 토큰화 처리
- Chroma Vector store: chroma DB에 토큰 정보를 저장
- Cosine similarity: 토큰 간의 유사도 연산

거대 언어모델 기반 대화형 건설안전교육 챗봇 프로토타입2



금일 작업 내용 알려줘

금일 작업 내용은 다음과 같습니다:

- 작업 사진을 찍은 내용은 내용콘크리트 양생이었습니다. 사진은 2장이었습니다.

그 이외의 작업 내용은 알 수 없습니다.

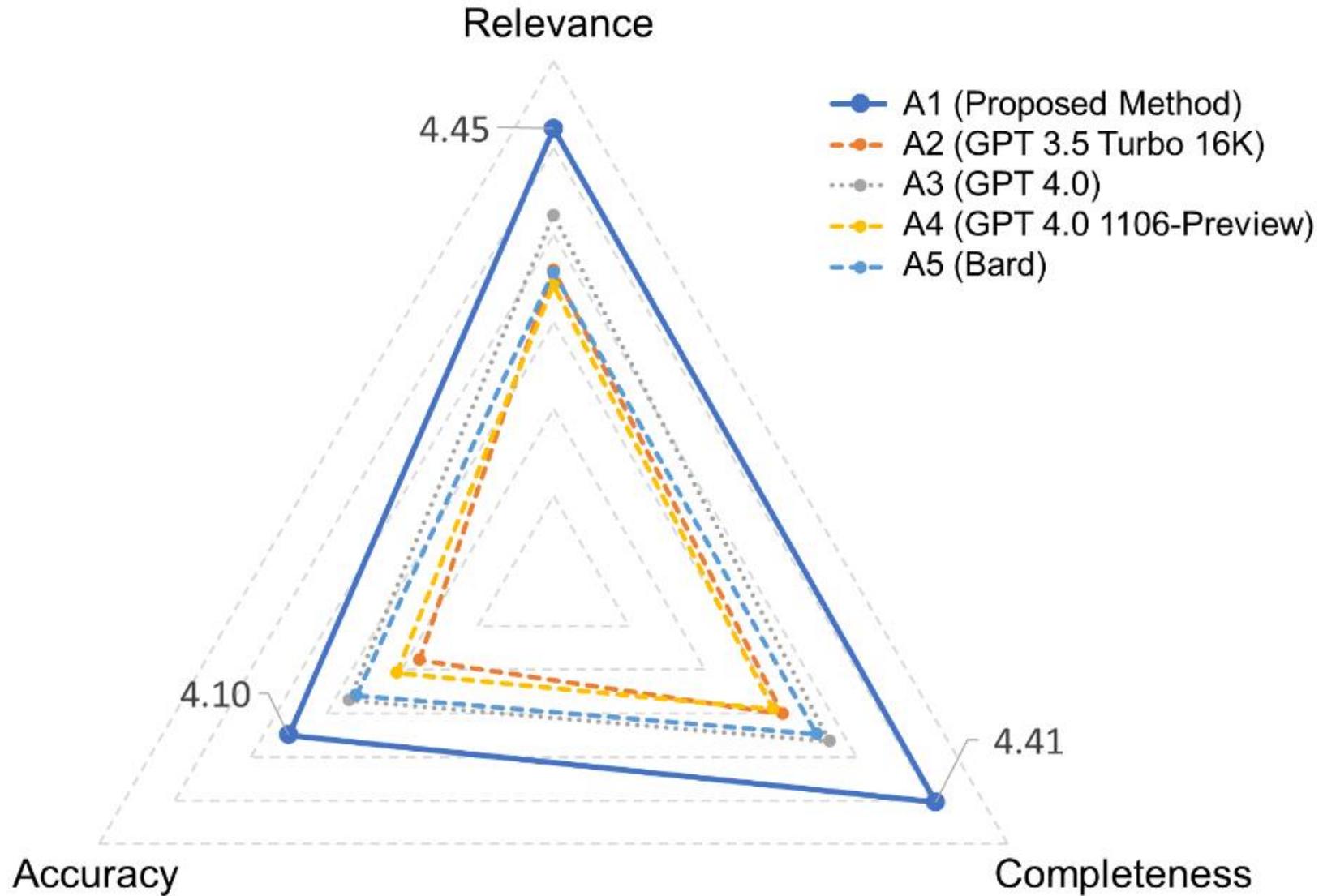
금일 작업 내용 알려줘

금일 작업 내용은 "비계 해체 정리"와 "수직 수평 가새재 해체하기" 그리고 "받침철물 해체하기"였습니다. 작업 내용에 대한 자세한 내용은 제공되지 않았습니다.

비계 해체 작업 절차 알려줘

비계 해체 작업의 절차는 다음과 같습니다: 1. 작업발판 발끝막이판 안전계단 경사로 해체하기 - 작업발판의 발끝막이판을 안전하게 해체합니다. 2. 벽연결용 철물 해체하기 - 비계와 벽을 연결하는 철물을 해체합니다. 3. 수직 수평 가새재 해체하기 - 비계의 수직과 수평 가새재를 해체합니다. 4. 받침철물 해체하기 - 비계의 받침철물을 해체합니다. 5. 자재 정리 정돈하기 - 해체한 자재를 정리하고 정돈합니다. 이렇게 순서대로 작업을 진행하면 비계 해체 작업을 완료할 수 있습니다.

평가 결과



결론

- 거대언어모델의 건설안전 분야 적용 방안 제시
- 자연어를 사용한 정보 검색 방법으로 건설 작업자의 정보 접근성 개선
- 기존의 건설안전 정보 통합에 소요되는 시간과 비용 절감
- 거대 언어 기반 생성형 AI에서 발생하는 허구나 편향된 정보 방지에 기여
- 건설 프로젝트 정보 및 개인 정보 노출 없이 작업자 및 작업 환경 맞춤형 건설안전 정보 제공의 가능성 제시
- 거대언어모델의 외국어 번역 기능 활용, 국내 외국인 노동자 대상의 서비스 확장 가능성 모색
- 대용량 데이터베이스와 거대언어모델과의 연결 및 응답 속도 개선 방안에 대한 연구가 진행 중

도로설계/공사 표준시방서 Chatbot

도로설계/공사 표준시방서 Chatbot

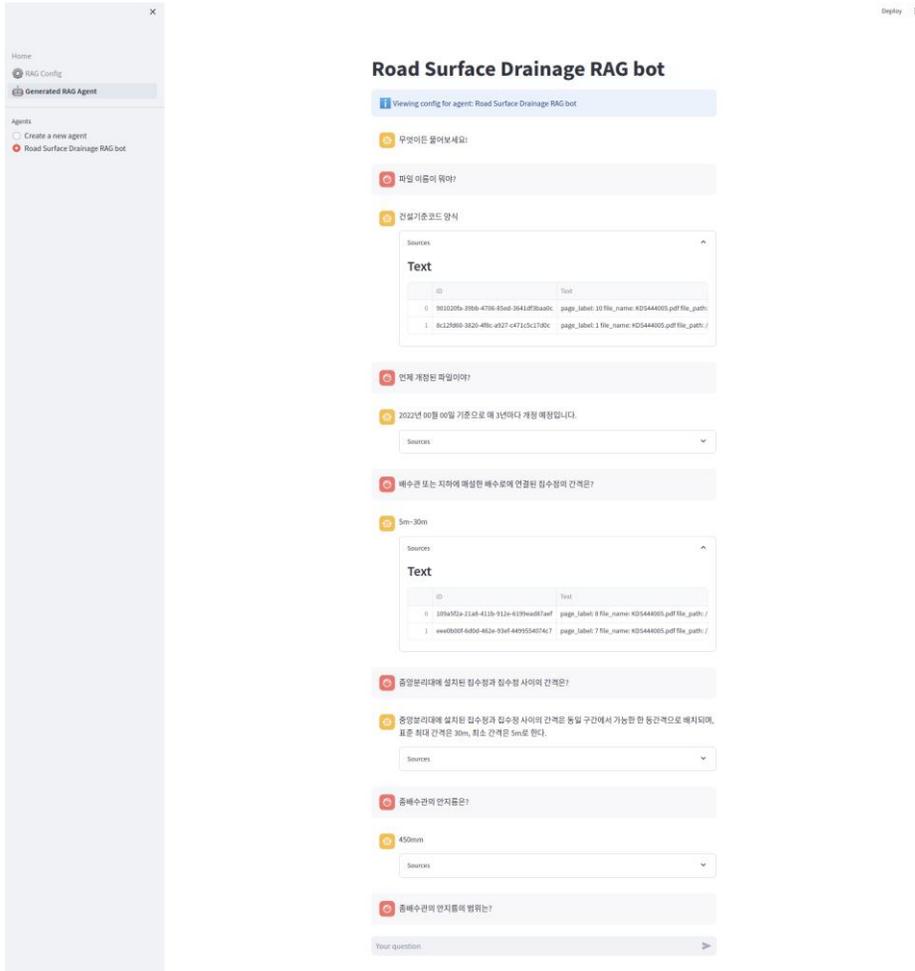
- 문제 정의:

브리콘 R&BD Factory 미팅 결과 기업이 보유한 구조계산서 등의 자료가 실제 설계 기준 및 표준시방서에 준하는지에 대한 판단 여부를 자동화하는 것에 대한 기업의 니즈가 있었음.

이에 RAG 기술을 적용하여 다양한 설계 기준에 대해 검토를 진행할 수 있는 Chatbot 시스템을 구현하였음.

→ 검토 과정에서 휴먼 에러를 최소화하여 정확성을 향상하고 시간을 절약하며 비용 절감 및 전반적인 프로젝트 관리를 개선하는데 기여할 수 있음.

도로설계/공사 표준시방서 Chatbot



- External Data
: 도로설계기준_노면배수.pdf
- Temperature = 0.1
- LLM = gpt-4-1106-preview(GPT 4 Turbo)
- Text Embedding Model = Text-embedding-ada-002-v2
- Top K = 2
- Chunk Size = 512
- Task
: Chatbot의 수식적/수치적 정확도 확인

QA Chatbot – 수치적 정확도

Chatbot

 종배수관의 규격은? 횡배수관과의 규격의 차이점은?

 종배수관의 규격은 450mm를 적용한다. 횡배수관은 포장층 내에 위치하는 경우 시멘트 콘크리트로 보강하고 최소 450mm 이상을 적용한다. 따라서, 종배수관과 횡배수관의 규격 차이점은 횡배수관이 보강이 필요한 경우가 있으며, 최소 규격이 450mm 이상이라는 점이다.

External DB

4.3.3 종배수관 및 횡배수관 최대길이

- (1) 중앙분리대 종배수관 규격은 450mm를 적용하며, 횡배수관은 포장층 내에 위치하는 경우 시멘트 콘크리트로 보강하고 최소 450mm 이상을 적용한다.
- (2) 종배수관의 최대길이는 청소 및 관 막힘을 고려하여 300~500m로 제한하여 횡배수 처리한다.

이외에도 도로설계기준_노면배수.pdf에 존재하는 모든 수치 정보에 대한 정확도는 100%를 달성함

OK! (4/4)

QA Chatbot – 수식적 정확도

Chatbot

 초기 집수정 설치 위치를 구하는 수식을 찾아줘.

 $Q=0.2778 \times 10^6 \times I \times (C_1 \cdot A_1 + C_2 \cdot A_2)$ (4.2-1)

다. 따라서 초기 집수정 위치는 다음과 같이 구할 수 있다.

$$Q=0.2778 \times 10^6 \times I \times (C_1 \cdot A_1 + C_2 \cdot A_2) \quad (4.2-1)$$

External DB

여기서, C_1 : 유출계수(포장부, 0.9)
 C_2 : 유출계수(절개부, 0.8)
 I : 강우강도 (mm/h)
 Q : 허용 통수량 (m^3/sec)
 A_1, A_2 : 유역면적 (m^2) (A_1 :포장부, A_2 :절개부)

 각각의 변수 및 계수 정보를 알려줘

 변수 및 계수 정보는 다음과 같습니다.

- C_1 : 유출계수 (포장부), 값은 0.9입니다.
- C_2 : 유출계수 (절개부), 값은 0.8입니다.
- I : 강우강도, 단위는 mm/h입니다.
- Q : 허용 통수량, 단위는 m^3/sec 입니다.
- A_1 : 유역면적 (포장부), 단위는 m^2 입니다.
- A_2 : 유역면적 (절개부), 단위는 m^2 입니다.

OK! (1/2)

QA Chatbot – 수식적 정확도

Chatbot



집수정 사이의 간격을 결정하는 수식을 찾아줘.



$S = C \times I \times W \times Q \times a$ (4.3-1)

② 집수정 간격(S)

가. 이 기준 4.2.1 (3)을 따라 식 (4.2-2)을 통해 산정한다.

$$S = \frac{3.6 \times 10^6 \times Q}{I(C_1 W_1 + C_2 W_2)} \quad (4.2-2)$$

External DB

여기서, Q : 허용 통수량 (m^3/sec)

S : 집수정 간격 (m)

I : 강우강도 (mm/h)

W_1, W_2 : 집수폭 (m)(W_1 :포장부, W_2 : 절개부)

C_1, C_2 : 유출계수(포장부, 0.9, 절개부, 0.8)

WRONG!(1/2)

QA Chatbot – 수식적 정확도

Chatbot

 C1과 C2가 2배가 될 때 집수정 간격(S)은?

 집수정 간격(S)은 식 (4.3-1)에 의해 계산되며, 여기서 C는 유출계수를 나타낸다. 해당 식에서 C의 값은 0.9로 주어져 있다.

만약 C1과 C2가 2배가 된다면, 각각 1.8과 1.6이 될 것이다. 그러나 집수정 간격(S)를 계산하는 식 (4.3-1)에서는 C 값만 사용되며, C1과 C2의 값은 직접적으로 S를 계산하는 데 사용되지 않는다. 따라서 C1과 C2가 2배가 되어도 집수정 간격(S)에는 영향을 주지 않는다. S는 여전히 원래의 계산방식대로 결정된다.

External DB

② 집수정 간격(S)

가. 이 기준 4.2.1 (3)을 따라 식 (4.2-2)을 통해 산정한다.

$$S = \frac{3.6 \times 10^6 \times Q}{I(C_1 W_1 + C_2 W_2)} \quad (4.2-2)$$

여기서, Q : 허용 통수량 (m^3/sec)

S : 집수정 간격 (m)

I : 강우강도 (mm/h)

W_1, W_2 : 집수폭 (m)(W_1 :포장부, W_2 : 절개부)

C_1, C_2 : 유출계수(포장부, 0.9, 절개부, 0.8)

WRONG!(1/2)

도로설계/공사 표준시방서 Chatbot

Challenges:

1. 시방서에는 문자 정보 외에도 다양한 표와 수식(그림) 정보가 존재하며, 이를 vector embedding 하는 과정에서 깨짐 현상 발생

비용·편익비 (B/C비율) = $\frac{\sum_n BPV_n}{\sum_m CPV_m} = TBPV / TCPV$ 

2. 구현된 Chatbot은 text 정보만 input으로 받고, external DB 역시 text 데이터만 vector화 가능

→ 기업이 보유한 구조계산서와 설계기준 검토 자동화가 불가능함

Equation Embeddings(2018)

arXiv:1803.09123v1 [stat.ML] 24 Mar 2018

Equation Embeddings

Kriste Krstovski¹ David M. Blei¹

Abstract

We present an unsupervised approach for discovering semantic representations of mathematical equations. Equations are challenging to analyze because each is unique, or nearly unique. Our method, which we call *equation embeddings*, finds good representations of equations by using the representations of their surrounding words. We used equation embeddings to analyze four collections of scientific articles from the arXiv, covering four computer science domains (NLP, IR, AI, and ML) and ~98.5k equations. Quantitatively, we found that equation embeddings provide better models when compared to existing word embedding approaches. Qualitatively, we found that equation embeddings provide coherent semantic representations of equations and can capture semantic similarity to other equations and to words.

1. Introduction

Equations are an important part of scientific articles, but many existing machine learning methods do not easily handle them. They are challenging to work with because each is unique or nearly unique; most equations occur only once. An automatic understanding of equations, however, would significantly benefit methods for analyzing scientific literature. Useful representations of equations can help draw connections between articles, improve retrieval of scientific texts, and help create tools for exploring and navigating scientific literature.

In this paper we propose equation embeddings (EqEmb), an unsupervised approach for learning distributed representations of equations. The idea is to treat the equation as a "singleton word," one that appears once but that appears in the context of other words. The surrounding text of the equation—and in particular, the distributed representations of that text—provides the data we need to develop a useful

representation of the equation.

Figure 1 illustrates our approach. On the left is an article snippet (Li et al., 2015). Highlighted in orange is an equation; in this example it represents a neural network layer. We note that this particular equation (in this form and with this notation) only occurs once in the collection of articles (from arXiv). The representations of the surrounding text, however, provide a meaningful context for the equation. Those words allow us to learn its embedding, specifically as a "word" which appears in the context of its surroundings. The resulting representation, when compared to other equations' representations and word representations, helps find both related equations and related words. These are illustrated on the right.

EqEmbs build on exponential family embeddings (Rudolph et al., 2016) to include equations as singleton observations and to model equation elements such as variables, symbols and operators. Exponential family embeddings, like all embedding methods, define a *context* of each word. In our initial EqEmb, the context for the words is a small window, such as four or eight words, but the context of an equation is a larger window, such as sixteen words. Using these two types of contexts together finds meaningful representations of words and equations. In the next EqEmb, which builds on the first, we consider equations to be sentences consisting of equation units, i.e., variables, symbols, and operators. Equation units help model equations across two types of context—over the surrounding units and over the surrounding words.

We studied EqEmbs on four collections of scientific articles from the arXiv, covering four computer science domains: natural language processing (NLP), information retrieval (IR), artificial intelligence (AI) and machine learning (ML). We found that EqEmbs provide more efficient modeling than existing word embedding methods. We further carried out an exploratory analysis of a large set of ~87k equations. We found that EqEmbs provide better models when compared to existing word embedding approaches. EqEmbs also provide coherent semantic representations of equations and can capture semantic similarity to other equations and to words.

¹Columbia University, New York, NY 10027. Correspondence to: Kriste Krstovski <kriste.krstovski@columbia.edu>, David M. Blei <david.blei@columbia.edu>.

Copyright 2018 by the author(s).

• Problem Statement

- 수식은 각각의 형식이 매우 다르고 또 복잡하기 때문에 다루기가 매우 까다로움.
- 따라서 수식의 이해를 자동화하는 것은 scientific literature를 탐색하는 데 유용할 것임.

• Hypothesis

- 수식 앞뒤의 텍스트 정보가 수식 자체의 유용한 표현을 생성하기 위한 데이터를 제공할 수 있다고 가설을 세움.

Equation Embeddings(2018)

Tensor Network RNTN (Socher et al., 2013) enables richer compositions between early time output e_{t-1} and **current word embedding** e_t . For **recurrent version**, let $e = [h_{t-1}, e_t]$. The **output embedding** at e_t is **given** by:

$$h_t = f(e^T \cdot W \cdot e + V \cdot e) \quad (6)$$

Similar method can be **extended** to **recursive model**. Let $e = [e_{\text{left}}, e_{\text{right}}]$ **denote** the **concatenation** of the two children's **vectors**. Parent representation e_η is given by:

$$h_t = f(e^T \cdot W \cdot e + V \cdot e)$$

Top Equations

1. $i_t = \sigma(W_{ix}x_t + W_{ir}r_{t-1} + W_{ic}c_{t-1} + b_i)$
2. $e_\eta = f(e^T \cdot W \cdot e + V \cdot e)$
3. $c_t = f_t \cdot c_{t-1} + i_t + l_t$
4. $h_t^s = o_t \cdot c_t$
5. $c_1 = \text{encode}_{\text{fixed}}(n, m, \text{start})$

Top Words

1. recurrent
2. input
3. layer
4. embedding
5. network



Equation Embeddings(2018)

- Experiment

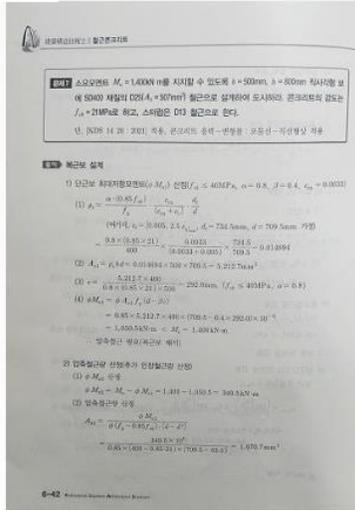
- arXiv에서 발표된 NLP, IR, AI, ML 분야 논문 87k(LaTeX format)
- 각 수식마다 2개의 단어를 샘플링하여 학습 데이터셋 생성
- 수식은 하나의 단어로 취급

- Results

$F\text{-measure} = \frac{2\text{PrecisionRecall}}{\text{Precision}+\text{Recall}}$		$\cos(\mathbf{x}, \mathbf{y}) = \frac{\sum_{i=1}^n x_i \cdot y_i}{\sqrt{\sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n y_i^2}}$					
Top Equations		Top Words	Rank	CBOW	PV-DM	GloVe	EqEmb
1.	$Recall = \frac{\text{Number of matched frames(color and texture)}}{\text{Number of frames in user summary}}$	1. f-measure	1.	vectors	vectors	accepted	cosine
2.	$ROUGE - N - P = \frac{\sum_{I \in CT} \sum_{gram_N \in I} C_{nt_match}(gram_N)}{\sum_{I \in CT} \sum_{gram_N \in I} C_{nt}(gram_N)}$	2. test	2.	inner	inner	brothers	distance
3.	$precision = \frac{n_{correct}}{n_{total}}$	3. report	3.	space	vector	partition	similarity
4.	$P@k = \frac{\text{number of actual friends at top } k}{k}$	4. performance	4.	center	angle	stimulus	metric
5.	$FNR = \frac{FN}{FN+TP}, FPR = \frac{FP}{FP+TP}$	5. accuracy	5.	good	dense	miss	cluster

Cross-sectional design

GL You



ChatGPT
The image you've uploaded contains mathematical formulas related to the calculation of the maximum resistance moment for reinforced concrete beams. Since you've requested the extraction of these formulas, I'll transcribe them from the image:

For the balanced section calculation (M_{n1}):

$$1. \rho_s = \frac{\alpha \cdot (0.85 f'_c)}{\beta \cdot f_y} \cdot \frac{e_{cu}}{(e_{cu} + e_y)} \cdot \frac{d}{d}$$

Given:

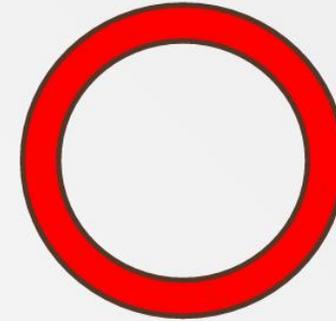
$$\alpha = 0.8, \beta = 0.4, e_{cu} = 0.0033$$

$$e_y = [0.005, 2.5 \frac{f_y}{E_s}], d' = 734.5mm, d = 709.5mm$$

- $A_{s1} = \rho_s \cdot b \cdot d = 0.014694 \cdot 500 \cdot 709.5 = 5,212.7mm^2$
- $c = \frac{0.8 \cdot (0.85 \cdot 21 \cdot 500)}{400} = 292.0mm$
- $\phi M_{n1} = \phi A_{s1} f_y (d - \beta c) = 0.85 \cdot 5,212.7 \cdot 400 \times (709.5 - 0.4 \cdot 292.0) \times 10^{-6}$

This results in:

$$\phi M_{n1} = 1,050.5kN \cdot m < M_u = 1,400kN \cdot m$$



```
Analysis
b = 500 # Width of the beam in mm
d = 709.5 # Effective depth in mm
d_prime = 734.5 # Depth to the location of the compression
fc_prime = 21 # Concrete compressive strength in MPa
phi = 0.85 # Strength reduction factor

# Calculation of balanced section (Mn1)
# Step 1: Calculate the strain in the steel at the balance
ey_balanced = min(0.005, 2.5 * fy / Es) # Steel strain at
# Step 2: Calculate rho_s (reinforcement ratio for balance
rho_s_balanced = (alpha * (0.85 * fc_prime)) / (beta * fy)
# Step 3: Calculate As1 (area of steel for balanced section
As1_balanced = rho_s_balanced * b * d
# Step 4: Calculate c (depth of the neutral axis for balance
c_balanced = (phi * (0.85 * fc_prime * b)) / (fy / (0.85 *
```

이미지
형식으로
표현된
Equation을
코드 형식으로
변환하여
해석하는 방법
적용 가능

Source: 서보현, 김태영. (2021.) 철근콘크리트(KDS 14 20, KDS 41 30). 예문사. 6-42.

04 발전 방향

수식 정보 이해도 향상 using LaTeX

4.3.2 집수정 간격 결정

(1) 집수정 설치간격

$$S = \frac{3.6 \times 10^6 \times Q \times a}{C \times I \times W} \tag{4.3-1}$$

여기서, S : 집수정 간격 (m)

C : 유출계수 (0.9)

I : 강우강도 (mm/h)

Q : 축대의 허용 통수량 (m^3/sec)

W : 집수폭 (m)

a : 보정계수

LaTeX format

PDF

```
\subsection*{4.3.2 집수정 간격 결정}

(1) 집수정 설치간격


$$S = \frac{3.6 \times 10^6 \times Q \times a}{C \times I \times W} \tag{4.3-1}$$


여기서,  $S$ : 집수정 간격 (m)

 $C$ : 유출계수 (0.9)

 $I$ : 강우강도 (mm/h)

 $Q$ : 축대의 허용 통수량 ( $\text{m}^3/\text{sec}$ )

 $W$ : 집수폭 (m)

 $a$ : 보정계수
```

04 발전 방향

수식 정보 이해도 향상 using LaTeX

집수정 사이의 간격을 결정하는 수식을 찾아줘.

Calling function: vector_tool with args: {"input": "What is the formula for determining the spacing between catch basins in road surface drainage?"}

집수정 사이의 간격을 결정하는 수식은 다음과 같습니다:

$$[S = \frac{3.6 \times 10^6 \times Q}{I \times (C_1 \times W_1 + C_2 \times W_2)}]$$

여기서:

- (S)는 집수정 사이의 간격(m)입니다.
- (Q)는 유출량(L/s)입니다.
- (I)는 강우 강도(mm/hr)입니다.
- (C₁)과 (C₂)는 각각 도로 표면의 유출 계수입니다.
- (W₁)과 (W₂)는 각각 도로 폭(m)입니다.

이 수식을 사용하여 도로 표면 배수 시스템에서 집수정 사이의 적절한 간격을 계산할 수 있습니다.



$$S = \frac{3.6 \times 10^6 \times Q \times a}{C \times I \times W}$$

구조계산 결과 검토 자동화

- ◆ 도로 설계기준 및 표준시방서 내용을 기반으로 구조계산 결과 검토 자동화
- ◆ 구조계산식이 최신 버전의 설계기준 및 시방서 내용에 부합하는지 확인
- ◆ 엑셀 형식의 복잡한 구조계산서를 생성형 AI가 이해할 수 있도록 변환

1.4 캔틸레버부 바닥판 설계

1.4.1 설계 단면

H1	0.220 m	L1	0.000 m
H2	0.280 m	L2	0.500 m
H3	0.000 m	L3	0.000 m
H4	0.000 m	L4	0.000 m
H5	0.220 m	L5	0.300 m
H6	0.080 m	L6	0.608 m
H7	0.400 m	L7	1.250 m
R	∞	L8	0.158 m
Y	40.0 km/h	L9	0.000 m

콘크리트 단위중량 25.0 kN/m³
 무근콘크리트 단위중량 23.5 kN/m³
 포장 단위중량 23.5 kN/m³
 DB 1차로 하중 (Pr) 72.0 kN/m²
 ※ 방호벽 면적 : 0.140 m²

1.4.2 단면력 산정

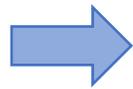
① 고정하중

NO.	산 식	작용거리(m)	모멘트(kN·m)	
1	0.000 × 25.000	0.000	1.158	0.000
2	½ × 0.000 × 0.000 × 25.000	0.000	0.908	0.000
3	0.500 × 0.000 × 25.000	0.000	1.158	0.000
4	½ × 0.000 × 0.000 × 25.000	0.000	0.908	0.000
5	0.500 × 0.280 × 25.000	3.500	1.158	4.051
6	1.408 × 0.220 × 25.000	7.741	0.704	5.448
7	½ × 1.250 × 0.000 × 25.000	0.000	0.574	0.000
8	0.158 × 0.000 × 25.000	0.000	0.079	0.000
9	0.908 × 0.080 × 23.500	1.706	0.454	0.774
10	1.000	1.000	1.158	0.000
TOTAL		13.947		11.431

② 활 하중 (DB-24 : Pr = 96 kN)

- 유효중 분포폭 (※ 도.설 3.6.1.4)
 $E = 0.5 \cdot X + 1.14 = 0.5 \times 0.608 + 1.14 = 1.63$ m 여기서, $X = 0.608$ m
- 충격계수
 $i = 15 / (40 + L) = 15 / (40 + 0.608) = 0.37 > 0.30$ $i = 0.300$
- 작용 모멘트
 $M^* = (72.00 / 1.626) \times 0.608 \times (1 + 0.30) = 34.97$ kN·m/m

5 페이지



1.4 캔틸레버부 바닥판 설계

1.4.1 설계 단면

H1	0.220 m	L1	0.000 m
H2	0.280 m	L2	0.500 m
H3	0.000 m	L3	0.000 m
H4	0.000 m	L4	0.000 m
H5	0.220 m	L5	0.300 m
H6	0.080 m	L6	0.608 m
H7	0.400 m	L7	1.250 m
R	∞	L8	0.158 m
Y	40.0 km/h	L9	0.000 m

콘크리트 단위중량 25.0 kN/m³
 무근콘크리트 단위중량 23.5 kN/m³
 포장 단위중량 23.5 kN/m³
 DB 1차로 하중 (Pr) 72.0 kN/m²
 ※ 방호벽 면적 : 0.140 m²

1.4.2 단면력 산정

① 고정하중

NO.	산 식	작용거리(m)	모멘트(kN·m)	
1	0.000 × 25.000	0.000	1.158	0.000
2	½ × 0.000 × 0.000 × 25.000	0.000	0.908	0.000
3	0.500 × 0.000 × 25.000	0.000	1.158	0.000
4	½ × 0.000 × 0.000 × 25.000	0.000	0.908	0.000
5	0.500 × 0.280 × 25.000	3.500	1.158	4.051
6	1.408 × 0.220 × 25.000	7.741	0.704	5.448
7	½ × 1.250 × 0.000 × 25.000	0.000	0.574	0.000
8	0.158 × 0.000 × 25.000	0.000	0.079	0.000
9	0.908 × 0.080 × 23.500	1.706	0.454	0.774
10	1.000	1.000	1.158	0.000
TOTAL		13.947		11.431

② 활 하중 (DB-24 : Pr = 96 kN)

- 유효중 분포폭 (※ 도.설 3.6.1.4)
 $E = 0.8 \cdot X + 1.14 = 0.8 \times 0.608 + 1.14 = 1.63$ m 여기서, $X = 0.608$ m
- 충격계수
 $i = 15 / (40 + L) = 15 / (40 + 0.608) = 0.37 > 0.30$ $i = 0.300$
- 작용 모멘트
 $M^* = (72.00 / 1.626) \times 0.608 \times (1 + 0.30) = 34.97$ kN·m/m

⑨ 피로에 대한 안전성 검토 (사용성 검토)

- 피로에 대한 안전성을 검토할 필요가 있는 경우, 철근의 응력이 다음의 값을 만족하면 피로에 대하여 검토할 필요가 없으나 초과하는 경우에는 적절한 방법으로 피로에 대하여 검토하여야 한다.

[피로를 고려하지 않아도 되는 철근의 응력범위 f_s^{max}]

철근의 종류	철근의 인장 및 압축응력범위
SD30	130 MPa
SD35	140 MPa
SD40	150 MPa

- $f_s = M_L / (A_s^{use} \times d \times j)$
 $= 29577600 / (1324.0 \times 180.0 \times 0.9118)$
 $= 136.1$ MPa < $f_s^{max} = 150.0$ MPa

0.KI



관련 내용에 해당하는 설계기준 파트에 접근하여 최신화 여부 확인

결론

- 검색증강기법(Retrieval Augmented Generation)을 활용한 거대언어모델은 환각(Hallucination) 현상을 획기적으로 감소
- 시방서 문서 기반 ChatBot 개발 가능
- 원본 파일 확보 필요
- Equation에 대한 정확한 해석 모듈 별도 개발 필요
- 건설 전문 문서 해석 LLM개발 후 설계 자동화 업무로 확장 가능성 확인

감사합니다.

김홍조

조교수

연세대학교 건설환경공학과

스마트 인프라 연구실

hongjo@yonsei.ac.kr

02-2123-2888

010-9658-1788

hongjo.github.io/