

캡스톤 디자인2 계획 발표

Scene Graph 지식 기반의 자연어 답변 생성 모델 연구

Natural Language Answer Generation Model Based on Scene Graph Knowledge

김민준, 송승우, 송지현



Contents

01 팀원 소개

02 연구 목표

03 모델 구축 계획

04 예상 결과물

05 차후계획



팀원 소개



김민준

역할 : VQA 논문 조사 및 연구



송승우

역할 : VQA 논문 조사 및 연구



송지현

역할 : 웹 개발

연구 목표

1. 캡스톤 1 모델의 확장 및 논문 작성
2. 학습한 모델을 웹을 통해 배포

연구 목표

- Visual Question Answering using Scene Graph
 - Scene Graph : 이미지 속 객체 간의 관계를 표현

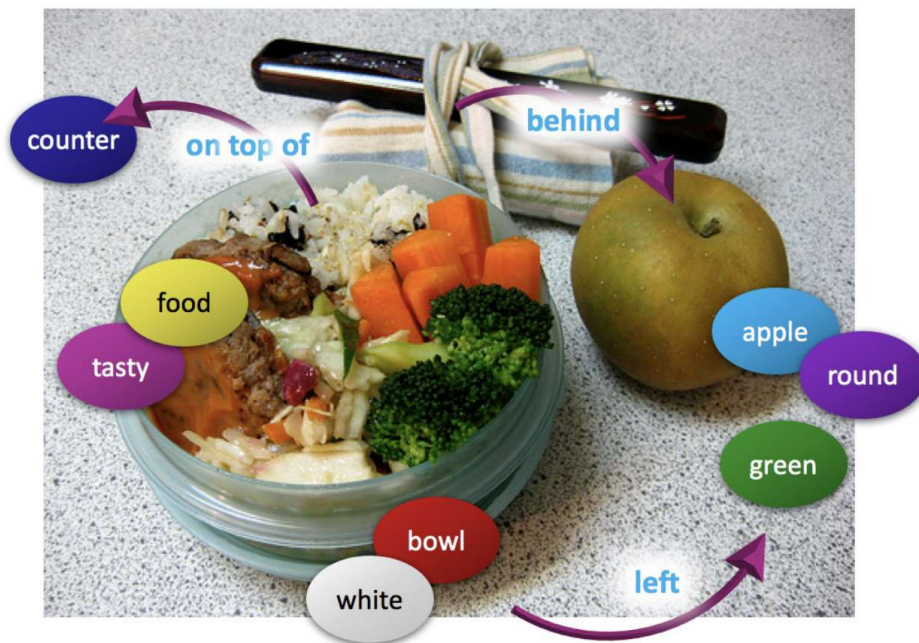


Figure 1: Examples from the new GQA dataset for visual reasoning and compositional question answering:

Is the *bowl* to the right of the *green apple*?

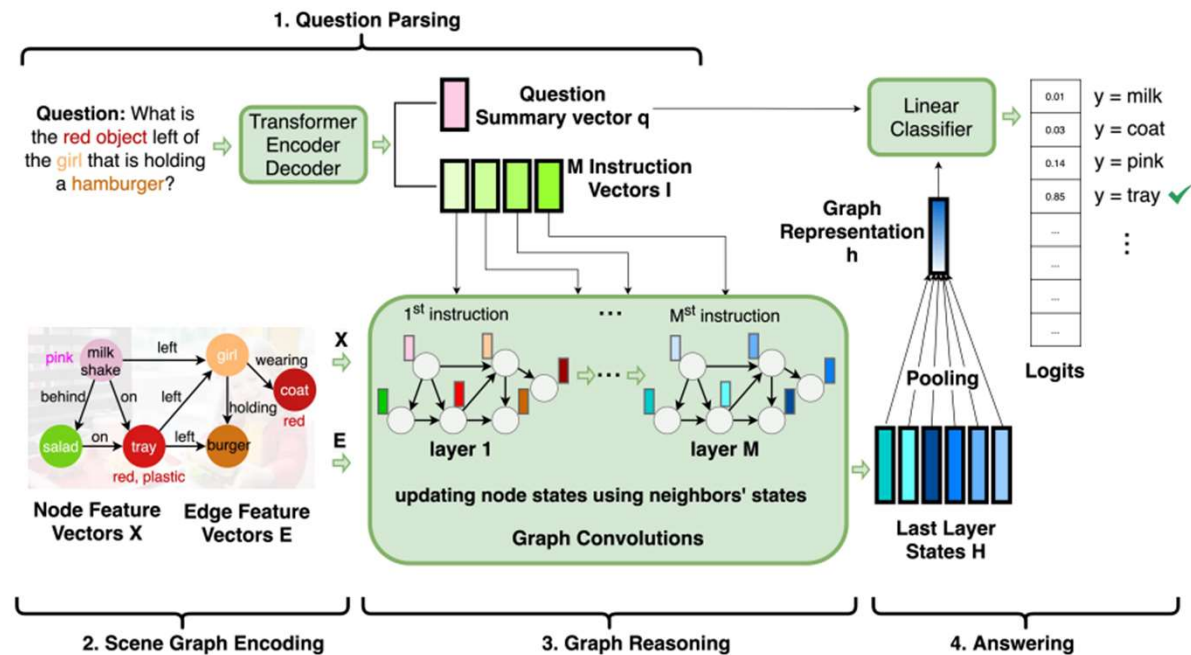
What type of *fruit* in the image is *round*?

What color is the *fruit* on the right side, red or *green*?

Is there any *milk* in the *bowl* to the left of the *apple*?

기존 연구

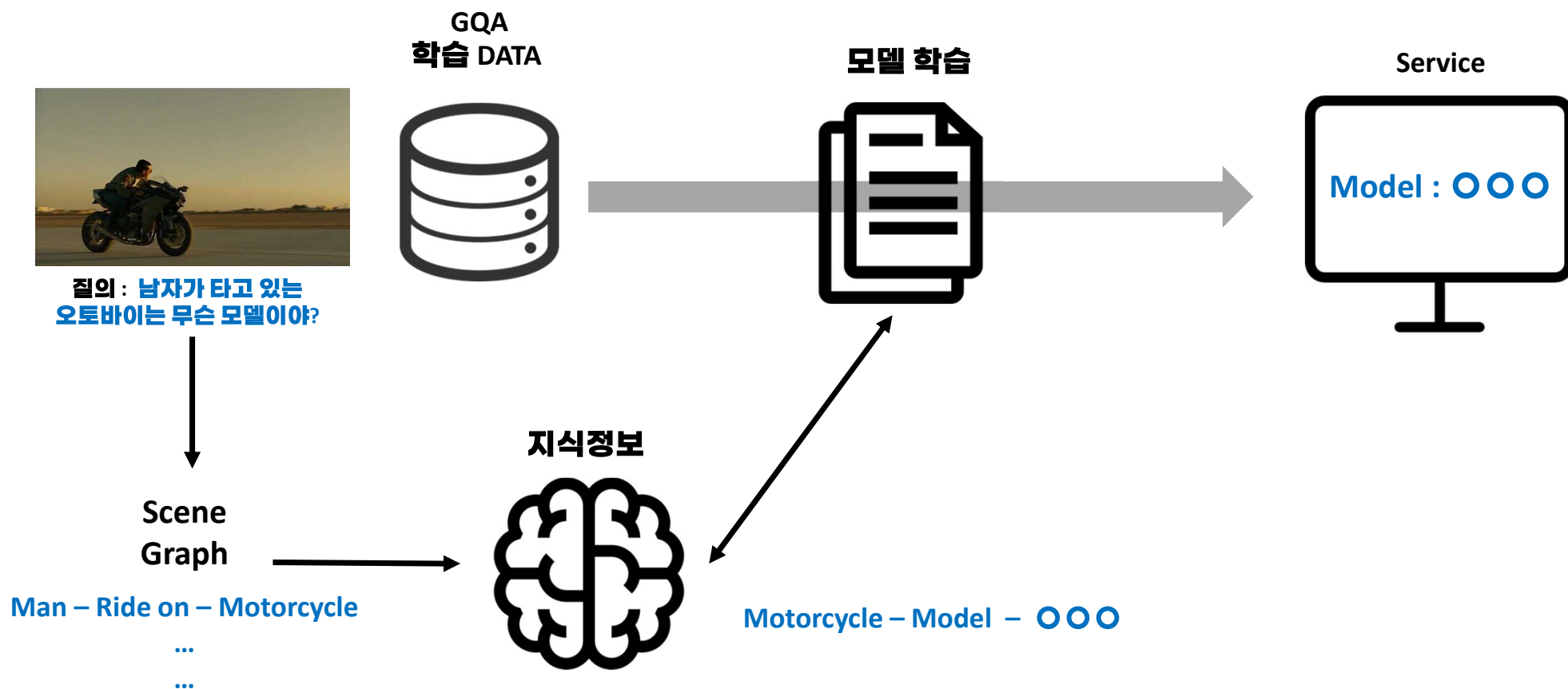
- GraphVQA: Language-Guided Graph Neural Networks for GQA



- ✓ Scene Graph을 사용하여 시각질의응답 모델 학습
- ✓ 한계점 : 분류를 통해 제한적인 답변을 도출

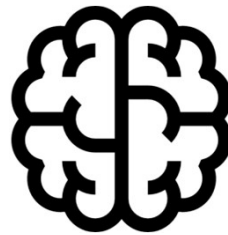
연구 목표

- Scene Graph & Knowledge Graph을 활용한 한국어 시각질의응답 모델 구축



연구 목표

- Scene Graph을 활용한 시각질의응답 모델 구축



Scene Graph & Knowledge Graph

- ✓ 캡스톤 1의 Knowledge Graph와 Scene Graph을 활용하여 시각질의응답 모델 구축

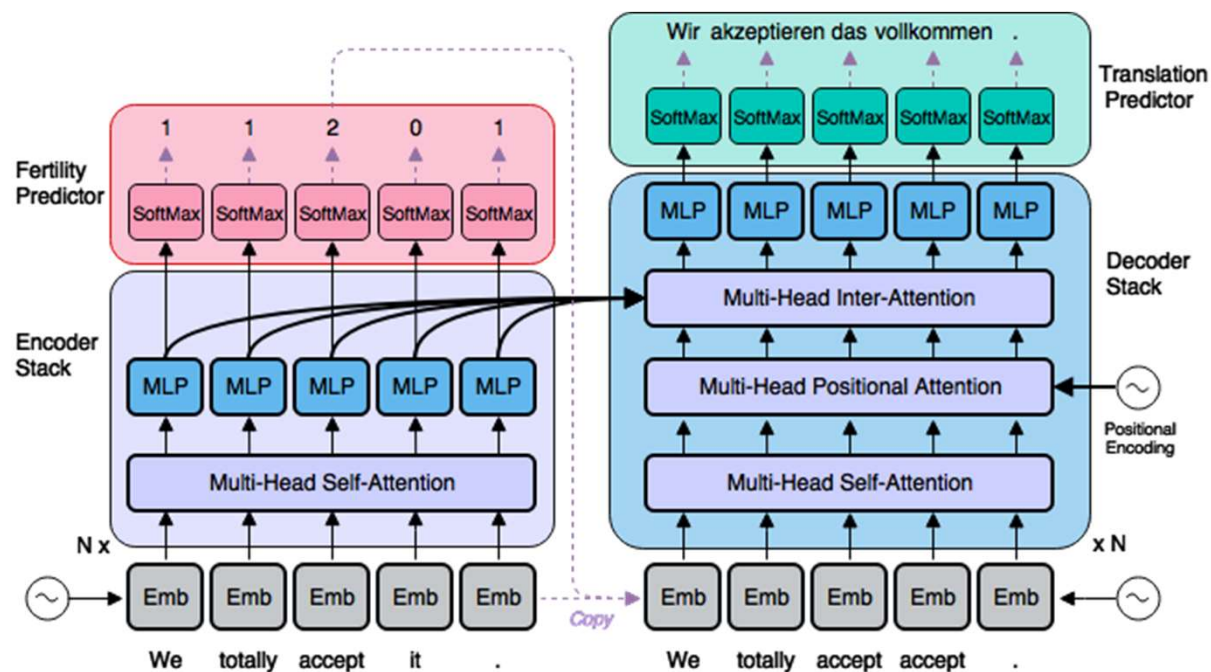
03 Scene Graph & GELVQA



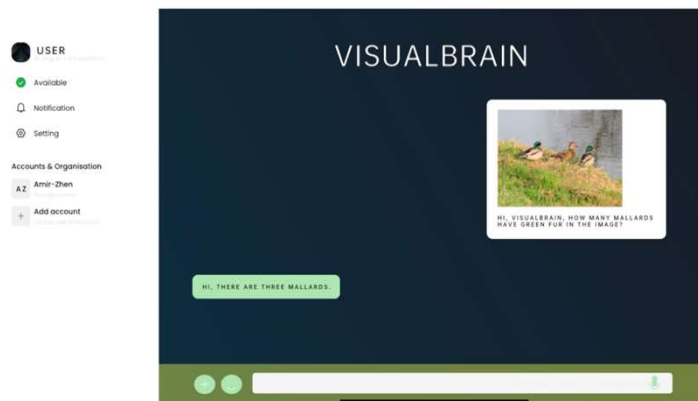
Scene Graph

+

Text Generation Model



04 예상 결과물



1. 연구 내용 논문 작성

- AAAI-2024 (08/15)
- Coling-2024 (10/13)

2. 모델 웹 배포

- (2주) : 학습된 PyTorch 모델 웹과 연동
- (2주) : 프론트엔드 개발, 사용자 인증, 챗봇 구현.
- (3주) : 백엔드 개발, 콘텐츠 제공, 진행 상황 추적 구현.
- (2주) : 테스트링, 버그 수정, 사용자 수용 테스트(UAT).
- (1주) : 배포 및 런칭.

05 차후계획

	7월	8월	9월	10월	11월	12월
GQA 자료조사						
모델 구축						
논문 작성						
웹 개발						