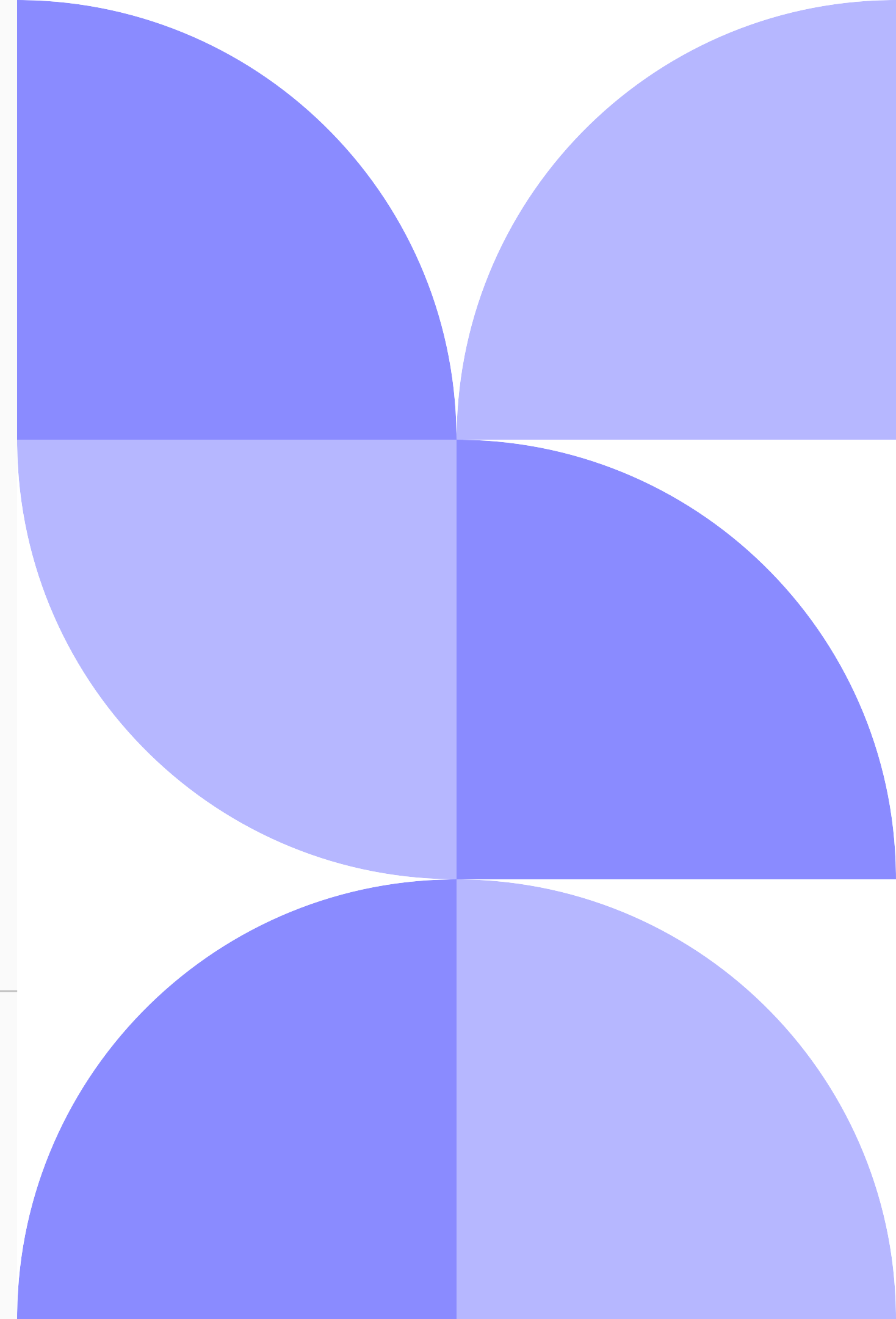


2025 2학기

VQA 스터디 중간 공유회

김태경 / 박찬욱



목차



1. 스터디 프로젝트 소개



2. 데이터셋 구축 방법



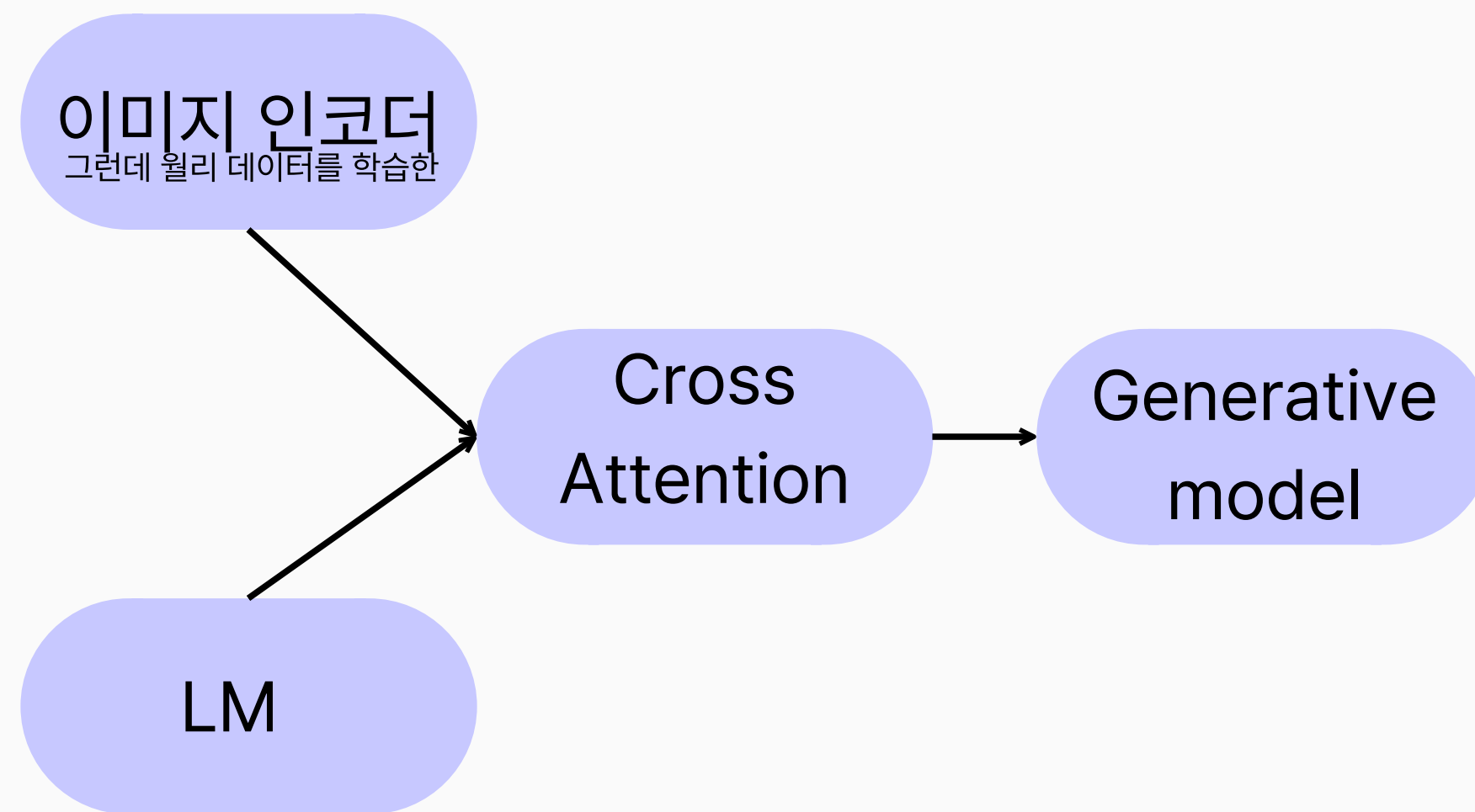
3. 향후 활동 안내

2학기 프로젝트



숨은그림을 찾고, 이에 대해 설명까지 해주는
만능 숨은그림찾기 VQA 모델

프로젝트 구상

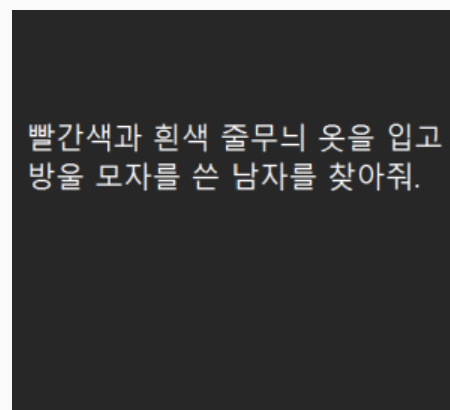


월리를 찾아라 전용 데이터셋으로 이미지 인코더를 학습/파인튜닝 시킴
언어모델과의 CROSS-ATTENTION으로 질문의 의미를 파악
월리와 이에 대한 설명을 생성 모델로 답변
이에 대해 진행이 잘 될 경우, 12월에 더 다양한 숨은그림찾기에 적용

데이터셋 구성

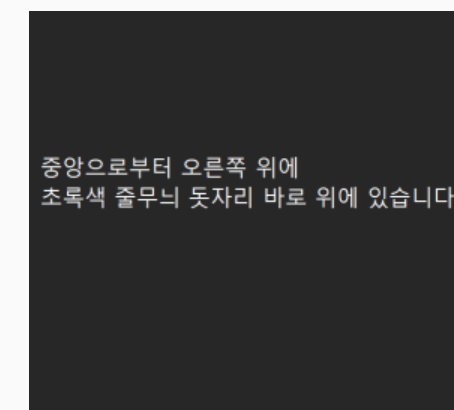
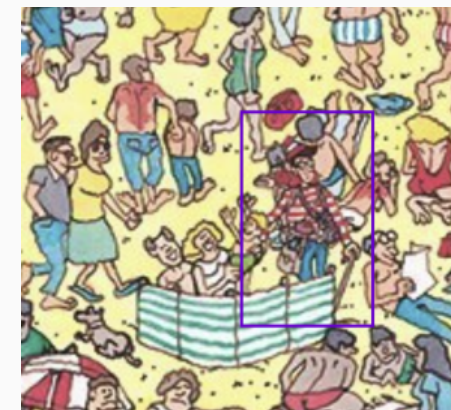
크게 데이터 수집 단계와, 데이터 어노테이션 단계로 구분

DATA COLLECTION



"월리를 찾아라" 책 구입 후 스캔을 통해 고화질의 이미지 데이터 확보 예정
이미지와 함께, 이미지에서 무엇을 찾을지에 대한 질문을 생성형 AI를 통해
각 이미지별로 30개씩 생성 예정

DATA ANNOTATION

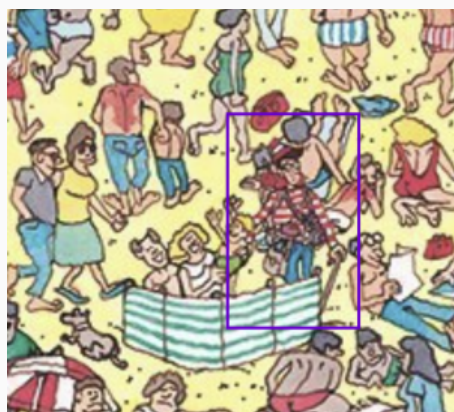


수집한 이미지와 질문을 사용 가능한 하나의 데이터로 가공하는 과정
바운딩박스로 질문에 대한 해답의 위치를 표시
해답의 위치에 대해, 근처를 자세히 묘사한 답변을 제시
최종적으로 수집하고 가공한 데이터를 전부 하나의 JSON 형식으로 구성

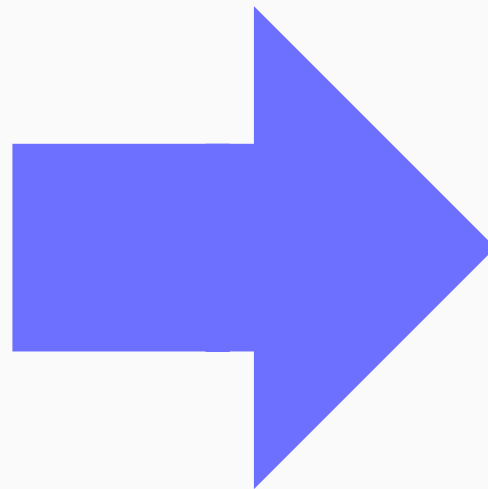
데이터셋 구성



빨간색과 흰색 줄무늬 옷을 입고
방울 모자를 쓴 남자를 찾아줘.



중앙으로부터 오른쪽 위에
초록색 줄무늬 돛자리 바로 위에 있습니다.



```
{  
  "IMAGE_ID": "WALLY_BEACH_001.JPG",  
  "IMAGE_PATH": "/DATA/IMAGES/WALLY_BEACH_001.JPG",  
  "WIDTH": 1920,  
  "HEIGHT": 1080,  
  "ANNOTATIONS": [  
    {  
      "QUERY_ID": "Q_001_01",  
      "QUESTION": "월리는 어디에 있나요?",  
      "TARGET_OBJECT": "월리",  
      "BOUNDING_BOX": [620, 450, 645, 490], // [X_MIN, Y_MIN, X_MAX, Y_MAX]  
      "ANSWER_DESCRIPTION": "이미지 중앙, 노란색 파라솔 아래에 있습니다."  
    }  
  ]  
}
```

향후 계획 설명



~ 11.10 - 데이터셋 수집



11.10 ~ 11.24 - 모델 학습



11.24 ~ 12.7 - 모델 완성 및 테스트



이후 - 월리 외 데이터로 재학습 및 적용

Q&A