

캡스톤디자인I 중간보고서

제 목	국문	시각적 추론 Generation VL 모델을 활용한 Multi-Modal ChatBot 어플리케이션		
	영문	Multi-Modal ChatBot Application using Generation VL Model with Visual Reasoning		
진 행 상 황	중요마일스톤	<p><주요 기능 : 모델 학습, 웹 서비스 구현></p> <ol style="list-style-type: none"> 1. 거대 시각적 추론 데이터셋으로 모델의 시각적 표현 학습 2. Image - Text fusion 모델 구축 및 학습 3. Generation type downstream task 모델 구축 및 학습 4. 학습이 완료된 모델로 웹 서비스 구현 		
	진행상황	<p>기존의 시각적 추론 데이터셋 조사 완료, 시각적 추론 관련 논문 조사 완료, PGM 기반 시각적 데이터셋 분석 완료, 시각적 추론 베이스라인 모델 생성 완료, SimCLR, SupCLR loss 모델에 적용 완료. 진행한 연구에 관해 2023 KCC 논문 제출 완료, 멀티모달 관련 논문 조사 중</p>		
산출물	요구사항 정의서(별첨 1), 중간보고서(별첨 2)			
팀 구성원	학년	학 번	이 름	연락처(전화번호/이메일)
	4	20181602	최창수	choics2623@gmail.com
	4	20181632	임현석	gustjrantk@gmail.com
	4	20201752	이현서	20201752@edu.hanbat.ac.kr
<p>컴퓨터공학과와 프로젝트 관리규정에 따라 다음과 같이 요구사항 정의서와 중간보고서를 제출합니다</p> <p style="text-align: center;">2023 년 04 월 27 일</p> <p style="text-align: right;">책임자 : (인) 지도교수 : (인)</p>				

[별첨1]

프로젝트명 : VL 어플리케이션 시스템 개발

소프트웨어 요구사항 정의서

Version 1.0

개발 팀원 명(팀리더):○○○
○○○
○○○
○○○
○○○

대표 연락처:000-000-000
e-mail: 000@000

목차

1. 개요
2. 시스템 장비 구성요구사항
3. 기능 요구사항
4. 성능 요구사항
5. 인터페이스 요구사항
6. 데이터 요구사항
7. 테스트 요구사항
8. 보안 요구사항
9. 품질 요구사항
10. 제약 사항
11. 프로젝트 관리 요구사항

1. 시스템 개요

사업의 목표: 멀티모달 질의응답 애플리케이션 개발을 통해 텍스트와 이미지를 모두 처리할 수 있는 인공지능 기반의 웹 서비스를 구축한다. 이 서비스는 VL(Vision Language) 생성 모델을 사용하여 사용자의 질문에 대한 답변을 텍스트 형태로 제공할 것이다.

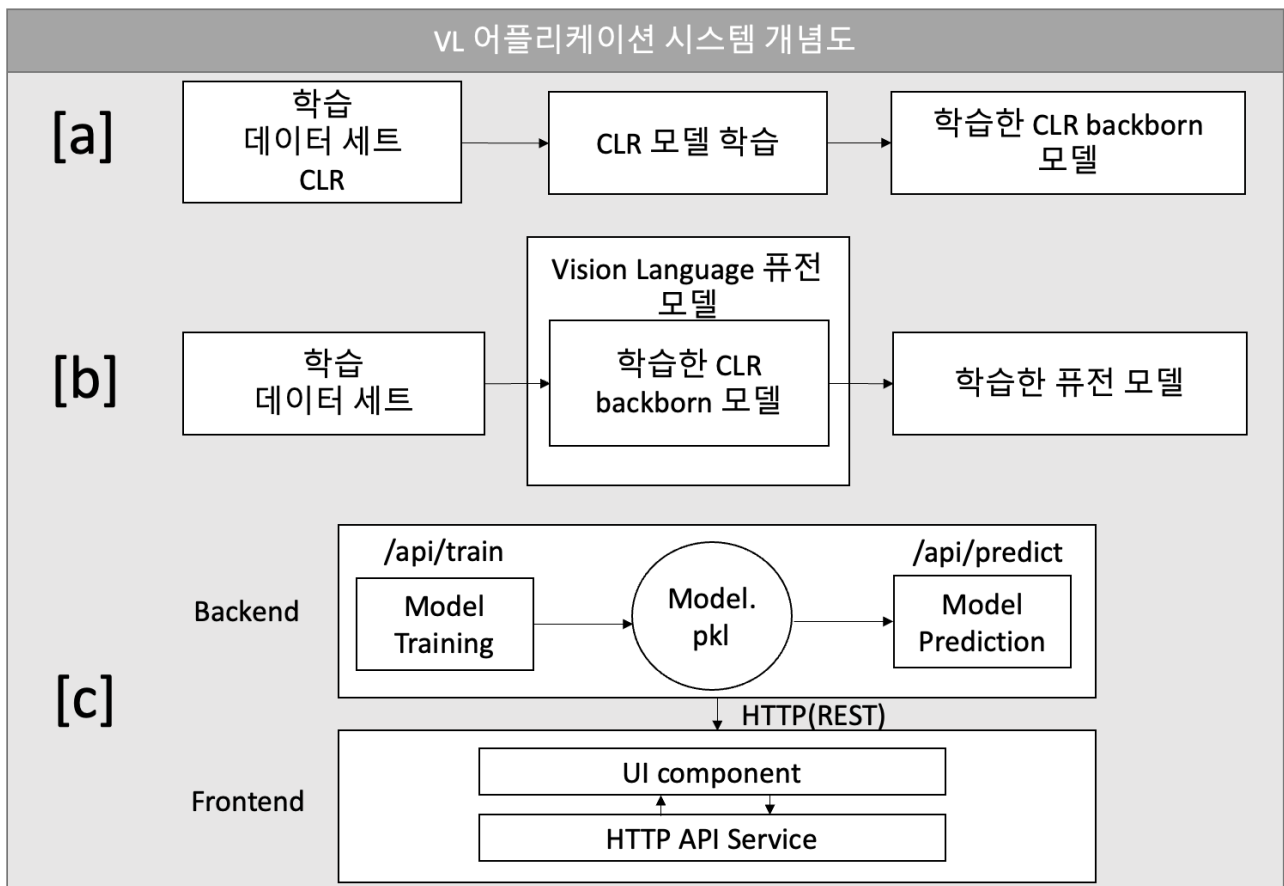
추진 범위:

1. 데이터 수집: 텍스트와 이미지 데이터를 모으고 정제하여 학습 데이터를 생성한다.
2. 모델 개발: 최신의 VL 생성 모델을 기반으로 새로운 멀티모달 질의응답 모델을 개발한다.
3. 웹 서비스 개발: 사용자가 이용할 수 있는 인터페이스를 제공하는 웹 서비스를 개발한다.
4. 모델 최적화 및 성능 평가: 개발된 모델의 성능을 평가하고 최적화하여 정확도를 높인다.
5. 지속적인 업데이트: 사용자의 요구와 시장의 변화에 맞춰 서비스를 지속적으로 업데이트한다.

시스템 구성도:

1. 사용자 인터페이스 (UI): 웹 페이지를 통해 사용자가 질문을 입력하고 결과를 확인할 수 있는 인터페이스를 제공한다.
2. 서버: 웹 서비스를 호스팅하며, 사용자의 요청을 처리하고 응답을 반환하는 역할을 담당한다.
3. 멀티모달 질의응답 모델: 서버에 배포되어 사용자의 질문을 처리하고, 텍스트와 이미지 형태의 답변을 생성한다.
4. 데이터베이스: 학습 데이터, 사용자 데이터, 서비스 관련 데이터를 저장하고 관리한다.

시스템 구성도



2. 시스템 장비 구성요구사항

요구사항 고유번호		ECR-001		
요구사항 명칭		장비 요구사항		
요구사항 분류		시스템 장비구성 요구사항	응락수준	필 수
요구사항 상세 설명	정의	모델 학습 장비		
	세부 내용	<ul style="list-style-type: none">- 장비 품목 : GPU (NVIDIA A100)- 장비 수량 : 4개- 장비 기능 : 모델의 연산 속도를 높인다.- 장비 성능 및 특징 : 메모리 80GB- 시간 제약사항 : 다른 캡스톤 팀과의 동일한 GPU 사용으로 할당 시간을 정해서 사용- 자원 제약사항 : 해당 사항 없음- 장애 처리 : 해당 사항 없음		

3. 기능 요구사항

요구사항 고유번호		SFR-001
요구사항 명칭		AI 모델 개발
요구사항 분류		기능
요구사항 상세 설명	정의	VL 어플리케이션 시스템 개발
	세부 내용	<ul style="list-style-type: none">• 대규모 시각 추론 데이터를 이용하여 이미지-텍스트 퓨전 모델에 비언어적 시각 표현을 학습 시킬 것• 기존의 거대 이미지-텍스트 데이터 셋을 퓨전 모델에 학습 시킬수 있도록 할 것• 클라우드 서버 시스템으로 사용자의 형태와 규모에 상관 없이 원활하게 서비스 운영이 가능하도록 할 것

4. 성능 요구사항

요구사항 고유번호		PER-001		
요구사항 명칭		처리 속도 및 시간		
요구사항 분류		성능 요구사항	응락수준	필 수
요구사항 상세 설명	정의	처리 속도 및 시간		
	세부 내용	모델이 이미지와 텍스트에 대해 답변하는 시간을 의미함		

5. 인터페이스 요구사항

요구사항고유번호		SIR001		
요구사항 명칭		대화창		
요구사항 분류		사용자 인터페이스	응락수준	필수
요구사항 세부내용	정의	대화창 구현		
	세부 내용	<ul style="list-style-type: none">- 대화창에는 사용자와 질의 응답 시스템 간의 대화가 표시될 수 있도록 인터페이스 구현- 사용자의 입력 창과 질의 응답 시스템의 출력 창을 구분하여 인터페이스에 표시		

6. 데이터 요구사항

요구사항고유번호	DAR-001		
요구사항 명칭	데이터 표준 준수		
요구사항 분류	데이터	응락수준	필수
요구사항 세부내용	- 데이터 전문 기관 유클라우드소프트 기관으로부터 검증받은 1.2TB 크기의 대규모 시각추론 학습 데이터 사용.		

요구사항고유번호	DAR-002		
요구사항 명칭	초기자료 구축		
요구사항 분류	데이터	응락수준	필수
요구사항 세부내용	- 올바른 CLR Loss 사용을 위한 초기 데이터 구축 및 사진과 json 파일로 관리		

7. 테스트 요구사항

요구사항고유번호	TER-001		
요구사항 명칭	성능 테스트		
요구사항 분류	테스트	응락수준	필수
요구사항 세부내용	<ul style="list-style-type: none"> - 구축된 모델이 원하는 sequence를 제대로 generation 하는 지를 테스트하고 점검하기 위한 평가 기준으로 loss와 acc로 평가 		

요구사항고유번호	TER-002		
요구사항 명칭	성능 테스트		
요구사항 분류	테스트	응락수준	필수
요구사항 세부내용	<ul style="list-style-type: none"> - CLR Information이 기존의 모델보다 성능이 높은 지 평가 - 이를 위한 평가 기준으로 기존 모델의 loss, acc와 CLR Information을 사용한 모델의 loss, acc를 비교 		

8. 보안 요구사항

요구사항고유번호	SER-001		
요구사항 명칭	보안지침 준수		
요구사항 분류	보안	응락수준	필수
요구사항 세부내용	- 사용자 데이터 세트를 배포한 BryanPlummer의 보안 및 저작권 관련 데이터 이용정책에 따라 개발이 수행되어야 함		

9. 품질 요구사항

요구사항고유번호	QUR-001		
요구사항 명칭	데이터 품질 관리		
요구사항 분류	품질	응락수준	필수
요구사항 세부내용	정의	품질 관리(기술 관점)	
	세부 내용	- 학습 데이터는 중복이 없어야 하며, 정제 과정에서 중복 데이터를 제외해야 한다. - 카테고리 별 데이터 수량은 균일하고 적절한 비율을 유지해야 한다. - 카테고리 라벨은 분류체계에 따라 명확하게 정의되어 있어야 하며, 각 카테고리 간 모호성이 없어야 한다. - 학습 데이터 및 검증, 테스트 데이터는 학습에 방해되지 않도록 각 이미지의 결측 정도 혹은 깨진 정도를 확인 후 제외하거나 포함시키는 판단을 거친 후 사용해야 한다.	

10. 제약 사항

요구사항고유번호	COR-001		
요구사항 명칭	시스템 개발과 설계 및 구현 제약사항		
요구사항 분류	제약사항	응락수준	필수
요구사항 세부내용	<ul style="list-style-type: none">- 현재 보유하여 활용 가능한 H/W, S/W를 최대한 활용함.- 대부분의 인공지능 모델 개발에 사용되는 Python(언어), PyTorch(프레임워크)를 사용함.- Sequence Generation에 좋은 성능을 보이는 GPT 모델을 기반으로 성능 도출		

11. 프로젝트 관리 요구사항

요구사항고유번호	PMR-001		
요구사항 명칭	품질관리		
요구사항 분류	프로젝트 관리	응답수준	필수
요구사항 세부내용	<p>세부 작업 분할 구조</p> <ol style="list-style-type: none"> 분석 <ul style="list-style-type: none"> - CLR Loss 중 SimCLR, SupCLR 사용 기법 조사 데이터 수집 및 생성 <ul style="list-style-type: none"> - VL 모델에 사용할 데이터 수집 시스템 설계 <ul style="list-style-type: none"> - 대규모 시각 추론 데이터를 이용한 CLR Loss 구현, GPT 모델 기반의 시각 질의 응답 시스템 개발 실험 <ul style="list-style-type: none"> - 대규모 시각 추론 데이터 셋을 이용한 CLR Information을 활용한 VL 모델을 사용하여 이미지와 해당 이미지를 설명하는 Flickr 데이터 셋을 활용하여 Sequence Generation 모델 생성. 해당 모델의 Flickr 데이터 셋에 대한 정확도 측정. - 프로젝트 수행조직에 대한 구성, 역할 <p>최창수 : 논문 및 자료조사, 모델 코드 작성 임현석 : 논문 및 자료조사, 모델 코드 작성 이현서 : PPT 및 자료조사, UI 개발</p>		

[별첨2]

중간보고서

1. 요구사항 정의서에 명시된 기능에 대하여 현재까지 분석, 설계, 구현(소스코드 작성) 및 테스트한 내용을 기술하시오.

Visual reasoning의 네 가지 카테고리에 대해 평가 모델을 구현하였으며, 현재 연구의 편의성을 고려하여 PyTorch Lightning 프레임워크로 코드를 변경하고 있다.

이 연구에서는 각 카테고리별 평가 모델의 성능을 철저히 분석하고, PyTorch Lightning을 활용하여 연구 과정의 효율성 높인다. 여러 카테고리에 대한 평가 모델을 구현하고 테스트한 결과를 통해, Visual reasoning 알고리즘의 각 카테고리별 성능 향상 방안을 제안하게 될 것이다. 최종적으로 도출된 모델로 다양한 Visual reasoning 문제에 대한 해결 능력을 평가하며, 해당 분야의 연구 발전에 기여할 수 있는 토대를 마련한다. 이러한 과정을 통해, Visual reasoning에 관한 국제 저널에 논문을 작성하고 제출할 계획이다.

2. 프로젝트 수행을 위해 적용된 추진전략, 수행 방법의 결과를 작성하고, 만일 적용과정에서 문제점이 도출되었다면 그 문제를 분석하고 해결방안을 기술하시오.

- 추진 전략 : 매주 담당 교수님에게 프로젝트 진행 사항을 발표하며, 교수님의 피드백과 팀원들과 토의함으로써 수월하게 진행함
- 수행 방법 : 각자 장비 환경에 맞추어 효율적으로 업무를 분담하고, 또한 '구글의 공유 플랫폼'을 적극적으로 사용하여 효율성을 높일 것
- 추진 전략 및 수행 방법의 결과 : 업무 효율이 높았으며, 역할 분담을 함으로써 모든 팀원의 참여도가 높았음
- 팀원의 책임 및 역할 수행에 대한 결과 : 업무를 균등하게 분담함으로써 원활하게 프로젝트 수행하여 큰 문제는 없었음
- 프로젝트 일정계획에 맞추지 못한 경우의 문제점과 해결 방안 : 모델 실험에 관련해서 환경 설정 문제에 봉착해 일정계획에 차질이 생겼지만, 원래 계획에 맞추기 위해 교수님과 추가 미팅함으로써 문제해결 방안을 알게 됐고 모든 팀원이 실험을 다시 진행하여 다음 일정계획에 차질이 없도록 함
- 요구사항 변경관리 : 변경 사항 없음

캡스톤 디자인 | 중간보고서 채점표

평가도구	평 가 항 목	평 가 점 수				
		1	2	3	4	5
중간 보고서 및 실행 결과	1. 요구사항 정의서(기능, 성능, 인터페이스 등)가 구체적으로 작성되었는가?					
	2. 요구분석, 설계 산출물(모델, 프로토타입 등)의 내용이 충실한가?					
	3. 설계 및 구현 문제를 위해 적용한 이론, 문제해결 방법이 제시되었으며 그 적용이 적합한가?					
	4. 구현된 소프트웨어(또는 이와 동등한 하드웨어 시스템)가 버그없이 실행되었는가?					
	5. 구현된 소프트웨어(또는 이와 동등한 하드웨어 시스템)의 성능 요구사항은 충족되었는가?					
도구활용	6. 설계 및 구현을 위해 도구가 적절히 활용되었는가?					
	7. 도구의 활용수준(능숙도)은 프로젝트 수행에 적합한가?					
팀원의 업무 및 역할	8. 팀원의 업무분담에 따른 역할 및 협력이 충실히 이루어졌는가? (평가자에 의한 질의)					
	9. 프로젝트 중간 진척상황에 대해 팀원이 충분히 인지하고 있는가?(평가자에 의한 질의)					
합계						
*검토 의견(최종완료 때까지 보완해야할 점에 대해 작성 요망)						
심사위원(소속):		(이름)			(인)	