

캡스톤디자인 중간보고서

제 목	국문	항공영상 기반의 제로샷 객체 탐지 연구		
	영문	Zero-shot Object Detection Based on Aerial Imagery		
진 행 상 황	중요마일스톤	<p><주요 기능: 영상 합성 기법 적용 및 데이터 보강></p> <ol style="list-style-type: none"> 1. 군용 객체 데이터셋 수집 및 분석 2. 드론 환경 조건 반영 영상 합성 기획 3. 합성 영상 생성 파이프라인 구현 4. 합성 영상 활용 성능 비교 실험 <p><주요 기능: 항공영상 기반 객체 탐지></p> <ol style="list-style-type: none"> 1. 드론 기반 군용 객체의 시각적 특징 분석 2. 기존 객체 탐지 모델 구조 및 특징 비교 분석 3. 모델 학습 및 최적화 실험 진행 4. 정량적 성능 평가 및 검증 5. 논문 제출 		
	진행상황	<p>군용 객체 데이터셋 수집 및 분석 완료, 드론 환경 조건 반영 영상 합성 기획 완료, 합성 영상 생성 파이프라인 구현 완료, 합성 영상 활용 성능 비교 실험 완료, 드론 기반 군용 객체의 시각적 특징 분석 완료, 기존 객체 탐지 모델 구조 및 특징 비교 분석 완료</p>		
산출물	요구사항 정의서(별첨 1), 중간보고서(별첨 2)			
팀 구성원	학년	학 번	이 름	연락처(전화번호/이메일)
	4	20201735	박우진	010-2360-2799 / 20201735@edu.hanbat.ac.kr
	4	20222019	김다빈	010-9855-2041 / 20222019@edu.hanbat.ac.kr
<p>컴퓨터공학과의 프로젝트 관리규정에 따라 다음과 같이 요구사항 정의서와 중간보고서를 제출합니다</p> <p style="text-align: center;">2025 년 5 월 2 일</p> <p style="text-align: right;">책임자 : 박우진 (인) 지도교수 : 장한얼 (인)</p>				

프로젝트명 : 항공영상 기반의 제로샷 객체 탐지 연구

소프트웨어 요구사항 정의서

Version 1.0

개발 팀원 명(팀리더): 박우진
김다빈

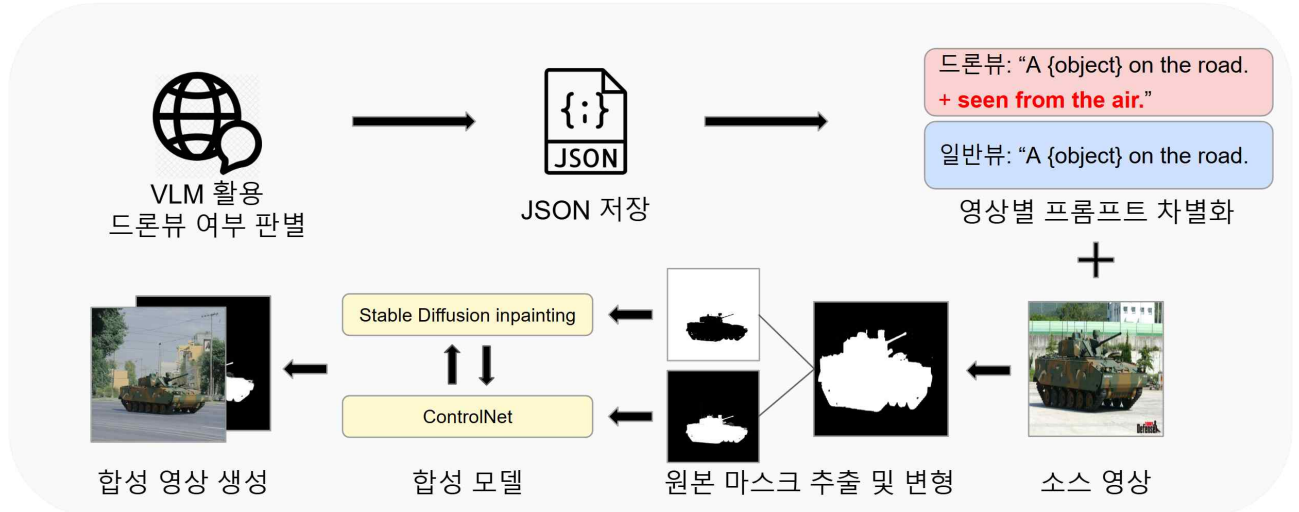
대표 연락처: 010-2360-2799
e-mail: 20201735@edu.hanbat.ac.kr

목 차

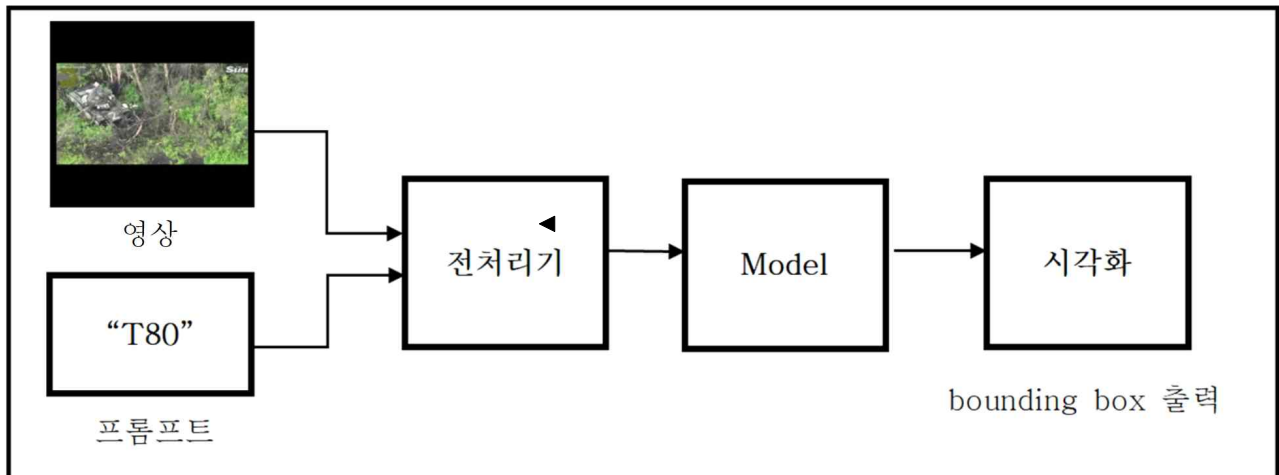
1. 개요	4
2. 시스템 장비 구성 요구사항	5
3. 기능 요구사항	5
4. 성능 요구사항	6
5. 인터페이스 요구사항	6
6. 데이터 요구사항	7
7. 테스트 요구사항	7
8. 보안 요구사항	7
9. 품질 요구사항	8
10. 제약 사항	8
11. 프로젝트 관리 요구사항	9
12. 중간보고서	10

1. 시스템 개요

- 드론 촬영 영상 및 실제 데이터를 입력으로 군용 객체 이미지를 생성하는 딥러닝 기반 이미지 합성 파이프라인 구조



- 합성 및 실제 영상 데이터를 활용해 텍스트 프롬프트와 멀티모달 인코딩을 통해 군용 객체를 탐지하는 자연어 조건부 제로샷 객체 탐지 시스템 구조



2. 시스템 장비 구성 요구사항

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 A6000) - 장비 수량 : 2개 - 장비 기능 : 모델의 연산 속도를 높인다 - 장비 성능 및 특징 : 개당 VRAM 48GB 		

3. 기능 요구사항

요구사항 고유번호		SFR-001		
요구사항 명칭		AI 모델 개발		
요구사항 분류		기능	응락수준	필 수
요구사항 상세 설명	정의	항공영상 기반의 제로샷 객체 탐지 모델 개발		
	세부 내용	<ul style="list-style-type: none"> - 드론 영상 기반의 제로샷 객체 탐지 모델 개발 - CLIP 기반 멀티모달 임베딩 구조 적용 - 자연어 프롬프트를 통한 미학습 객체 탐지 기능 구현 - 다양한 환경 대응을 위한 합성 데이터 활용 및 실험 수행 		

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	<ul style="list-style-type: none"> - 드론 영상을 입력으로 받아 객체 탐지 결과를 출력하기까지의 시간을 의미함 		

5. 인터페이스 요구사항

요구사항 고유번호		SIR-001		
요구사항 명칭		결과 시각화 인터페이스		
요구사항 분류		사용자 인터페이스	응락수준	필 수
요구사항 상세 설명	정의	탐지 결과를 확인할 수 있는 시각적 인터페이스 구현		
	세부 내용	<ul style="list-style-type: none"> - 탐지된 객체의 위치, 클래스 이름을 포함한 bounding box를 화면 상에 시각적으로 표시할 수 있는 인터페이스 구현 - 영상 입력 창과 모델의 예측 결과 출력 창을 구분하여 인터페이스에 표시 		

6. 데이터 요구사항

요구사항 고유번호		DAR-001		
요구사항 명칭		전처리된 데이터		
요구사항 분류		데이터	응락수준	필 수
요구사항 상세 설명	세부 내용	<ul style="list-style-type: none"> - 드론 영상 기반 군용 객체 이미지 데이터 수집 후 COCO JSON 형식으로 저장 - 영상 합성 기법을 통해 생성된 데이터 또한 동일한 구조로 변환하여 저장 		

7. 테스트 요구사항

요구사항 고유번호		TER-001		
요구사항 명칭		성능 테스트		
요구사항 분류		테스트	응락수준	필 수
요구사항 상세 설명	세부 내용	<ul style="list-style-type: none"> - mAP 정량 지표를 사용하여 제로샷 객체 탐지 모델의 성능 비교를 통한 탐지 정확도 평가 		

8. 보안 요구사항

요구사항 고유번호		SER-001		
요구사항 명칭		보안지침 준수		
요구사항 분류		보안	응락수준	필 수
요구사항 상세 설명	세부 내용	<ul style="list-style-type: none"> - 자체 수집한 드론 영상 데이터와 영상 합성 데이터의 보안 및 저작권 관련 데이터 이용정책에 따라 개발이 수행되어야 함 		

9. 품질 요구사항

요구사항 고유번호		QUR-001		
요구사항 명칭		데이터 품질관리		
요구사항 분류		품질	응락수준	필 수
요구사항 상세 설명	정의	품질관리(기술 관점)		
	세부 내용	<ul style="list-style-type: none"> - 드론 영상 기반의 실제 영상과 영상 합성 데이터 간의 비율이 적절히 구성되어야 함 - 학습 데이터는 중복이 없어야 함 - 학습 데이터 및 검증, 테스트 데이터는 학습 및 검증 결과에 방해되지 않도록 영상에 누락된 부분이 없는지 검토해야 함 - 객체가 명확하게 식별 가능한 품질의 영상으로 구성되어 있는지 검토해야 함 		

10. 제약 사항

요구사항 고유번호		COR-001		
요구사항 명칭		시스템 개발과 설계 및 구현 제약사항		
요구사항 분류		제약사항	응락수준	필 수
요구사항 상세 설명	세부 내용	<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 		

11. 프로젝트 관리 요구사항

요구사항 고유번호		PMR-001		
요구사항 명칭		프로젝트 관리		
요구사항 분류		프로젝트 관리	응락수준	필 수
요구사항 상세 설명	세부 내용	<ul style="list-style-type: none"> - 세부 작업 분할 구조: 1. 분석 <ul style="list-style-type: none"> - 드론 영상 기반의 객체 탐지 기술 및 제로샷 탐지 기법 조사 - 자연어 기반 프롬프트와 멀티모달 모델 활용 사례 분석 2. 데이터 수집 및 전처리 <ul style="list-style-type: none"> - 드론 촬영 실제 영상 수집 및 영상 합성 기반 데이터 생성 3. 시스템 설계 <ul style="list-style-type: none"> - Grounding DINO, YOLO-World 등 OVD 모델 구조 설계 및 비교 4. 실험 <ul style="list-style-type: none"> - mAP 정량 지표를 사용하여 제로샷 객체 탐지 모델 간 비교 실험을 통한 성능 측정 - 프로젝트 수행조직에 대한 구성, 역할: 박우진 : 데이터 수집 및 전처리, OVD 모델 백본 선정 및 최적화, 실험 결과 정리 및 논문 작성 김다빈 : 데이터 수집 및 전처리, OVD 모델 백본 선정 및 최적화, 실험 결과 정리 및 논문 작성 		

중간보고서

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

- 군용 객체 데이터셋 수집 후 데이터 확인



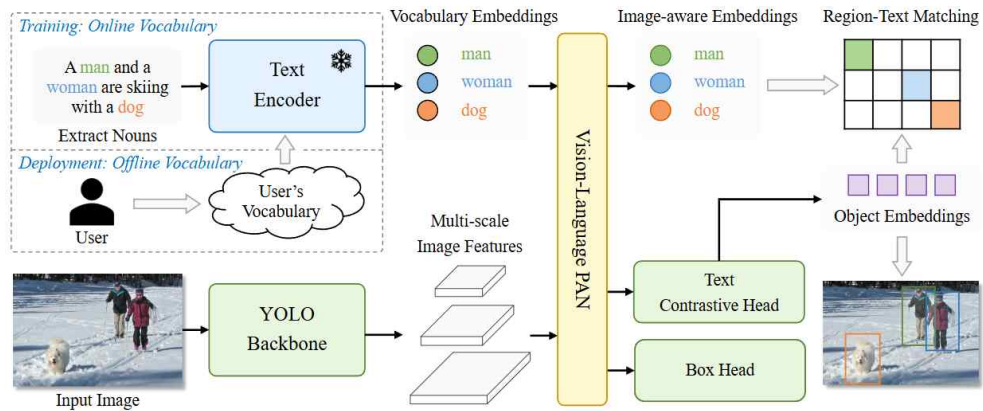
- 영상 합성 기법을 이용한 데이터 확인



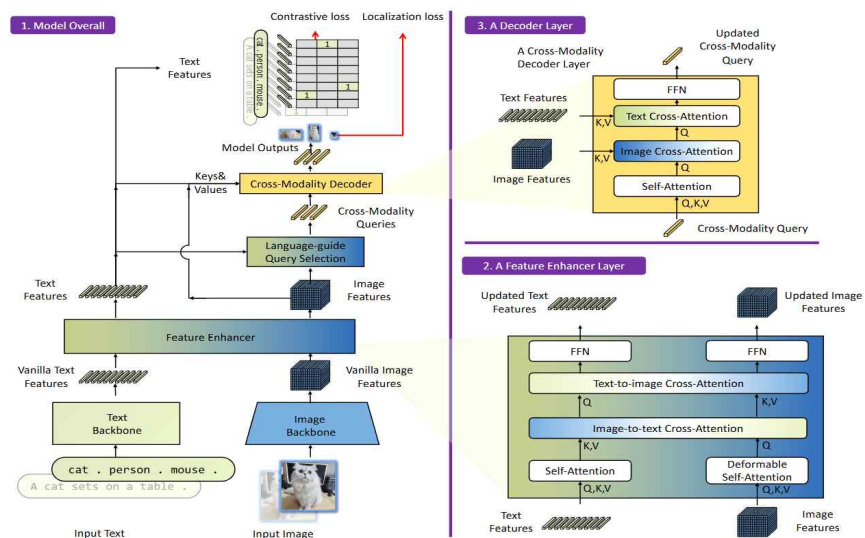
- 드론 기반 군용 객체의 시각적 특징 분석
 → 드론 기반 군용 객체의 시각적 특징 분석을 위해 Wikipedia에서 군용 객체 및 장비의 시각적 묘사에 사용되는 프롬프트를 조사하여 분석에 반영

- 기존 객체 탐지 모델 구조 및 특징 비교 분석

	논문 제목	학술지 / 학술대회	발행년도	탐지 방식	특징
1	YOLO-World: Real-Time Open-Vocabulary Object Detection	CVPR	2024	Zero-shot	<ol style="list-style-type: none"> 1. CSPDarknet 기반 경량 backbone을 사용하여 고속 feature 추출 2. Cross-attention 모듈을 네크와 헤드에 삽입하여 임베딩 수준에서 클래스 조건을 융합
2	Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection	ECCV	2024	Zero-shot	<ol style="list-style-type: none"> 1. Swin Transformer 기반 vision backbone으로 high-resolution feature 표현을 강화 2. 사전 학습된 BERT/CLIP 텍스트 인코더로 자연어 프롬프트를 처리 3. 디코더에 cross-attention 레이어를 추가해 텍스트 쿼리와 비전 쿼리를 효과적으로 융합



- 제로샷 객체 탐지 모델의 구조 1(YOLO-World)



- 제로샷 객체 탐지 모델의 구조 2(Grounding DINO)

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

- 추진 전략: 드론 영상 기반의 군용 객체 데이터셋은 보안 및 접근성 문제로 실제 촬영 영상 확보가 매우 제한적이었다. 이를 극복하기 위해 현실감 높은 합성 영상 생성 기법을 도입하여 부족한 데이터를 보강하였다. 구체적으로는 Stable Diffusion 모델과 ControlNet 모델을 활용하여 딥러닝 합성 기법을 통해 전차·장갑차·수송차 등의 외형·질감·배치가 실제와 최대한 유사하도록 자동으로 영상을 생성하였다. 이렇게 확보한 합성 데이터는 실제 드론 영상과 혼합 학습함으로써 다양한 작전 환경에서도 안정적인 탐지 성능을 확보하는 것을 목표로 하였다.

- 수행 방법: 각자 장비 환경에 맞추어 효율적으로 업무를 분담하고 구글의 공유 드라이브를 적극적으로 사용하였다.

- 프로젝트 일정 계획 지연 및 해결 방안: 군용 객체에 대한 제로샷 객체 탐지 연구과정에서 OVD(Open-Vocabulary Object Detection) 모델이 세부 분류 탐지 성능에서 한계를 보여 성능 확보가 어려울 것으로 보인다. 이로 인해 프로젝트 일정이 지연될 우려가 있다. 따라서 OVD 모델로 성능 확보가 불가능하다고 판단되는 시점에 VLM 계열 모델을 백업 플랜으로 선정하고, VLM 기반 병렬 실험을 실시하여 성능 개선 가능성을 검증·적용할 예정이다.

- 요구사항 변경관리의 결과: 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 버그 없이 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)	