

# 착수 보고서

Drone 영상에서의 Detection 및 Free-Space Segmentation



제출일	2023. 05. 22	팀명	FlyToTheSky
담당교수	박진선	팀장	김경현
		팀원	유일해, 신민건

## 목차

1. 과제 배경 .....	3
2. 목적 .....	3
3. 선행 연구 분석 .....	4
4. 진행 방안 .....	4
4.1 모델 구성 .....	4
4.2 데이터 구성 .....	5
4.3 시각화 방안 .....	6
5. 예상 문제점 .....	6
5.1 사용가능한 영역의 기준 .....	6
5.2 real-time segmentation의 구현 .....	6
5.3 데이터의 부족 .....	7
6. 개발 일정 및 역할 분담 .....	7
6.1 개발 일정 .....	7
6.2 역할 분담 .....	8
7. 참고 자료 .....	8

## 1. 과제 배경

드론으로 찍은 영상과 이를 활용해 Object를 detection 하거나 tracking 하는 것은 다양한 산업 분야에 유용하게 활용될 수 있다. 우리는 드론에 부착된 카메라로 촬영한 고화질 이미지를 분석하여 야외 적치장 내 화물 사용률 관리 시스템을 만들고자 한다. 이를 통해 물류 관리를 효율적으로 진행하고, 인건비를 줄일 수 있기 때문에 경제적 이익과 생산성 향상을 도모할 뿐만 아니라, 산업에서 일어날 수 있는 사고의 예방과 같은 부대효과도 기대할 수 있다.

## 2. 목적



그림 1. 평택항 부두의 야외 적치장

그림1에서 알 수 있듯, 야외 적치장이란 기업에서 대형 부품, 컨테이너 등을 수용하기 위해 사용되는 야외 창고 같은 개념이다. 컨테이너와 대형 부품들은 사람이 한 눈으로 규모를 예측할 수 없고, 야외 적치장의 규모가 큰 경우 해당 적치장 내의 물품이 얼마나 적재되어 있는지 아는 것은 어려운 문제이다. 따라서 본 과제의 일차적인 목표는 드론으로 촬영한 영상을 이용해 야적장의 사용률을 계산하는 것이다. 그 후에 넓은 야외 적치장의 사용률을 실시간으로 관리, 감독할 수 있다. 이를 위해 촬영된 영상을 가공하고 딥러닝 모델의 입력으로 넣어 Semantic segmentation을 수행한다. Semantic segmentation의 결과로 "비어 있는 공간", "적재된 공간", "적치장 이외의 공간" 등으로 class를 pixel 단위로 분류할 수 있고, 이미 적재된 공간과 더 적재할 수 있는 공간의 pixel의 개수를 이용해 현재 야외 적치장의 사용률을 계산할 수 있다.

이를 통해 기업에서 해당 야외 적치장의 사용률을 실시간으로 확인하여 비용, 시작적으로 효율적인 적치장 사용을 도모할 수 있다.

### 3. 선행 연구 분석

드론에 부착된 카메라를 이용하여 얻어진 이미지는 사람이 보는 시야가 제공하는 이미지와는 큰 차이가 있다. 이에 대해 [2] 논문에서는 natural image를 통해 pretrain된 IMP(ImageNet Pretrained) 형식의 모델들은 전이학습에 적합하지 않을 수 있다고 말하고 있다. 따라서 RSP(Remote Sensing Pretrained) 모델을 이용한 semantic segmentation 성능을 분석하여 IMP 방식과 RSP 방식을 비교한다. [3] iSAID Dataset에 대한 테스트 결과로 UperNet method를 이용한 IMP-viTAEv2-S, IMP-Swin-T 두 모델이 높은 mIOU 수치를 나타낸다. 일부 Class는 RSP에서 더 높게, 나머지는 IMP에서 더 높게 측정되었다. RSP에서 높은 지표를 나타낸 Class는 Baseball diamond, ground track field, Bridge, swimming pool 과 같이 비교적 큰 object들의 Class였다. 위 논문에서는 UperNet method를 사용한 RSP deep model은 high-resolution feature가 적절히 encoding되지 못했다고 판단하며 small object에 대한 segmentation은 향상될 필요가 있지만 특정 Class에 대해 매우 좋은 성능을 보인다고 언급하고 있다.

### 4. 진행 방안

#### 4.1 모델 구성

우리는 Upstream semantic segmentation task에서 높은 성능을 보여주는 IMP, RSP backbone model을 선별하여 전이학습을 통해 문제를 해결하려고 한다. 이를 위해 다양한 평가 지표 간의 차이를 분석할 필요가 있다. 우리는 정확도, mIOU, Itou, mean F1, Precision, Recall 등을 고려하여 모델의 성능을 평가하고 가장 적합한 backbone model을 선별할 예정이다. 선정된 model들은 우리가 해결하고자 하는 도메인의 데이터로 전이학습을 진행하고 이를 통해 모델이 특정 도메인의 특징을 더 잘 이해하고, 높은 성능을 발휘 할 수 있도록 개선 후 여러 실험을 진행할 것이다.

현재 나와 있는 여러 SOTA 모델들의 성능을 확인하고, 이후 실험 성능 결과를 바탕으로 논문 리뷰 및 실험을 진행하여 모델을 개선하고자 한다. 전이학습 - 논문리뷰 - 면담 후 좀 더 개선할 점을 찾고자 한다.

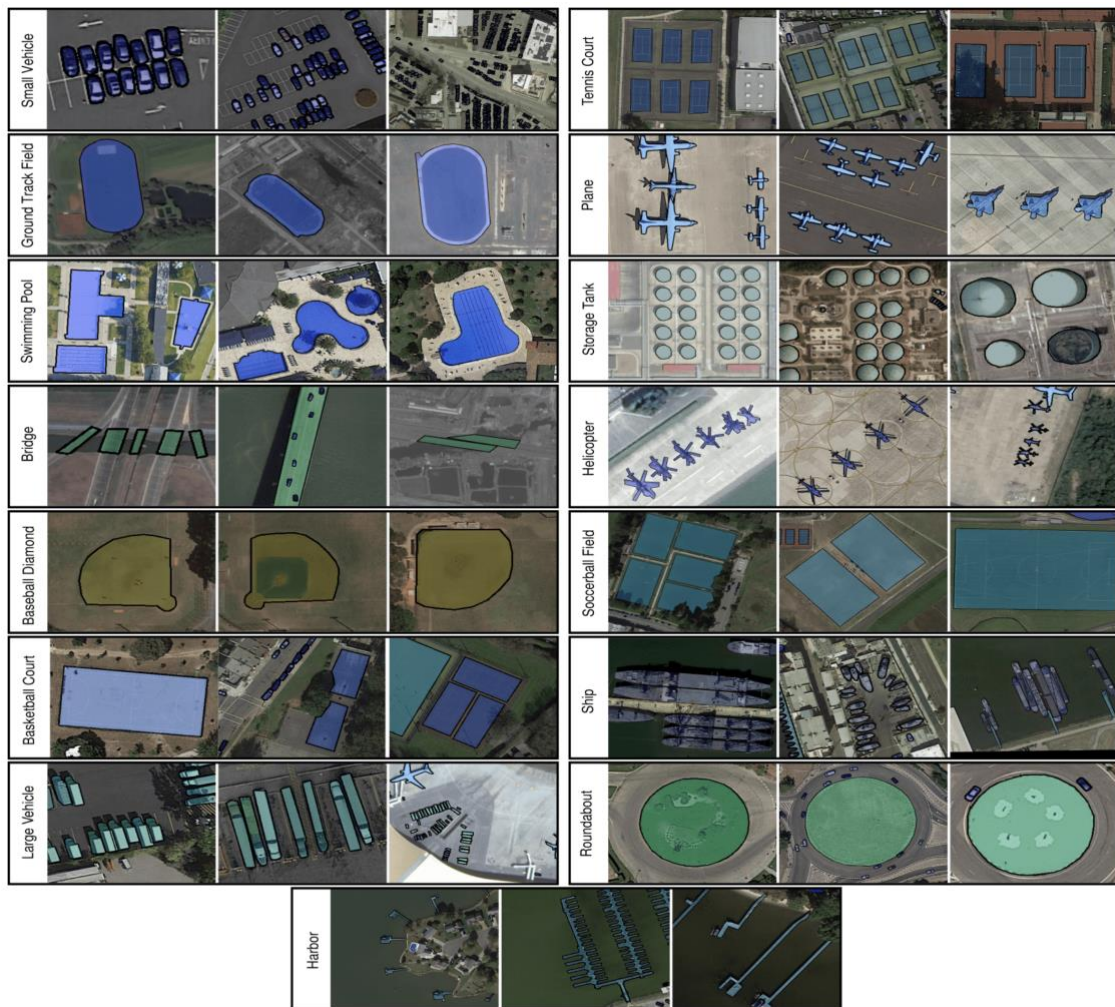


그림 2. label of iSAID Dataset

## 4.2 데이터 구성

본 과제에 사용할 dataset은 드론으로 촬영한 야외 적치장 사진이다. 다양한 종류의 컨테이너와 대형 부품들이 적치되어 있다. 하지만 이미지의 사이즈가 약 4000\*3000으로 매우 크고, 양이 적기 때문에 data augmentation하여 적절한 크기와, 증강이 필요하다. 해당 데이터셋은 현재까지 총 3개의 야외 적치장에 관한 영상이고, 데이터의 양이 제한적이기 때문에 앞에서 언급한 [3] iSAID 데이터셋을 사용하여 사전 학습한 모델을 추가로 학습시키는 방법으로 더욱 높은 정확도를 기대할 수 있다. iSAID 데이터셋은 구글 어스를 이용하여 얻은 이미지 데이터로, 일반적인 드론 촬영 고도보다는 다소 높은 곳에서 촬영되었지만, 1028\*1028 크기로 crop 한다면 비슷한 고도에서 촬영한 이미지처럼 활용할 수 있다. 해당 데이터셋에는 그림2와 같이 비행기, 테니스 코트, 농구 코트 등 총 15개의 label로 이루어져 있다.

### 4.3 시각화 방안

드론으로 촬영한 적치장 이미지는 적치장의 부분적 정보만을 제공한다. image matching algorithm과 드론의 위치 정보를 이용하여 이미지들을 합성하여 적치장 지도를 만들고 모델이 부분 이미지를 통해 만든 mask를 투영하여 적치장 사용 상태 지도를 시각화하고 사용율을 계산해 web을 통해 제공하고자 한다. 드론의 이동은 선형적이기 때문에 연속으로 촬영된 이미지에 대해 depth estimation과 같은 3D reconstruction 과정을 통해 사용자에게 더 직관적인 정보를 제공하는 방안도 시도해 볼 수 있다. 그러나 드론의 선형적 이동은 전체 정보에 대한 실시간 업데이트에 적합하지 않다. 부분적 정보의 실시간 업데이트, 전체 정보의 주기적 업데이트와 같은 정보 업데이트 논리 차이에 대한 결과 변화 논의와 더불어 전체 정보에 대한 빠른 업데이트 방법에 대한 논의가 필요하다. 또한 실제로 이용자가 드론으로 찍은 데이터를 업로드 하여 시각화한 결과를 볼 수 있는 웹페이지를 구현해 확인할 수 있도록 한다.

## 5. 예상 문제점

### 5.1 사용가능한 영역의 기준

야외 적치장은 그림1 에서 볼 수 있듯 컨테이너 혹은 대형 부품을 넣는 공간이 아파트 주차장처럼 규격화 되어 있지 않다. 따라서 어디서부터 어디 까지를 적재가능한 free space라고 규정할 것인지가 애매하다. 또한 free space는 차량이 드나드는 길과 큰 차이가 없기 때문에 차도와 free space 간의 기준을 명확하게 세워야 할 것으로 보인다.

### 5.2 real-time segmentation의 구현

실시간으로 segmentation이 이루어 지기 위해서는 가벼운 모델이 필요하다. object detection의 경우에도 one-stage detector가 아닌 two-stage detector 로는 실시간으로 object를 detection 하기 어려운 문제가 있다. 이와 비슷한 문제가 더 고차원의 문제인 segmentation에도 발생할 것이다. YOLACT: Real-time Instance Segmentation에서는 localization을 생략하고 업무를 병렬적으로 해결하는 방법으로 real-time을 구현하고 있다. 해당 부분을 논의 후 차용할 예정이다.

## 5.3 데이터의 부족

야외 적치장 사진은 150장 정도가 있다. 현실 문제를 해결하기 위해 프로젝트의 문제와 가장 비슷한 iSAID dataset의 SOTA모델을 사용할 것이다. 충분한 학습을 위해 데이터 증강 후 사용하려 한다. 실험의 결과가 나오지 않는다면 구글 earth를 사용해서 추가 데이터를 확보하거나 극단적으로는 물건이 적재되어 있는 곳과 아닌 곳을 swap하는 방법도 고려하여 데이터를 증강하고자 한다.

## 6. 개발 일정 및 역할 분담

### 6.1 개발 일정

6월			7월					8월					9월			
3주	4주	5주	1주	2주	3주	4주	5주	1주	2주	3주	4주	5주	1주	2주	3주	4주
segmentation 스터디																
	기존 모델 iSAID 적용															
				야외 적치장 데이터셋 모델 성능 확인												
						최적 모델 확인										
						중간 보고서 작성										
								모델 최적화								
											github 연동					
													최종 보고서 준비			
															포스터 제작	

## 6.2 역할 분담

이름	역할 분담
김경현	<ul style="list-style-type: none"><li>• 데이터 활용 방법 조사</li><li>• 모델 개발 및 평가</li><li>• 프로젝트 매니징</li></ul>
유일해	<ul style="list-style-type: none"><li>• 선행 연구 조사</li><li>• 데이터 관리 및 전처리</li><li>• 모델 개발</li></ul>
신민건	<ul style="list-style-type: none"><li>• 도메인 분석</li><li>• 데이터 전처리</li><li>• 모델 개발</li></ul>
공동	<ul style="list-style-type: none"><li>• 데이터 및 과제 분석</li><li>• github 관리</li><li>• 보고서 작성 및 포스터 제작</li><li>• 발표 및 시연 준비</li></ul>

## 7. 참고 자료

[1] <https://www.kyeonggi.com/article/201803220976964>

[2] An Empirical Study of Remote Sensing Pretraining (Di Wang, Jing Zhang, Bo Du, Gui-Song Xia, Dacheng Tao)

[3] <https://captain-whu.github.io/iSAID/index.html>

[4] YOLACT: Real-time Instance Segmentation (Daniel Bolya, Chong Zhou, Fanyi Xiao, Yong Jae Lee)