SW중심대학 공동 AI 경진대회 2023

동고동락

여은동 윤정윤 이다원 정재욱 조현성

프로세스

THERE ARE FIVE STAGES.

STAGE 1
데이터 전처리

STAGE 2
데이터증강

STAGE 3
모델

STAGE 4
후처리

STAGE 5
앙상블

STEP 1

데이터 전처리

PREPROCESSING

데이터 정보



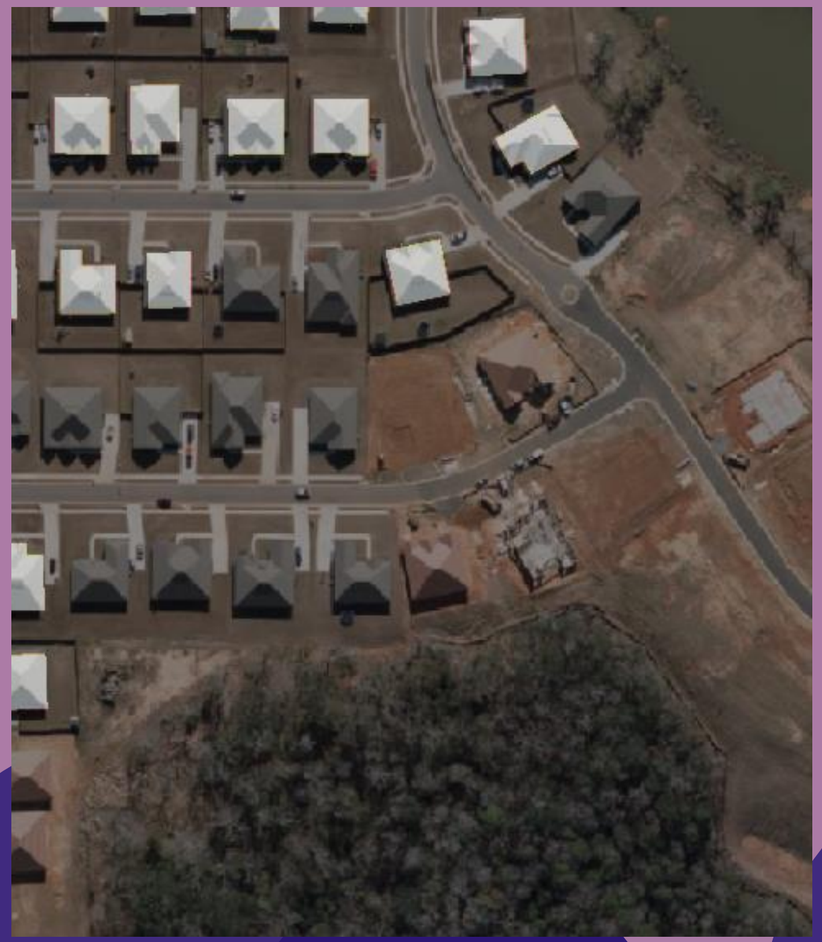



위성사진

- 대부분의 사진에 밀집되거나, 흩어져있는 숲 존재
- 같은 building이라도 여러 형태가 있음(큰 건물, 특이 건물)
- 이외 도로, 황무지, 호수 존재

PREPROCESSING

학습데이터 취사선택



잘못된 마스킹 이미지 제거

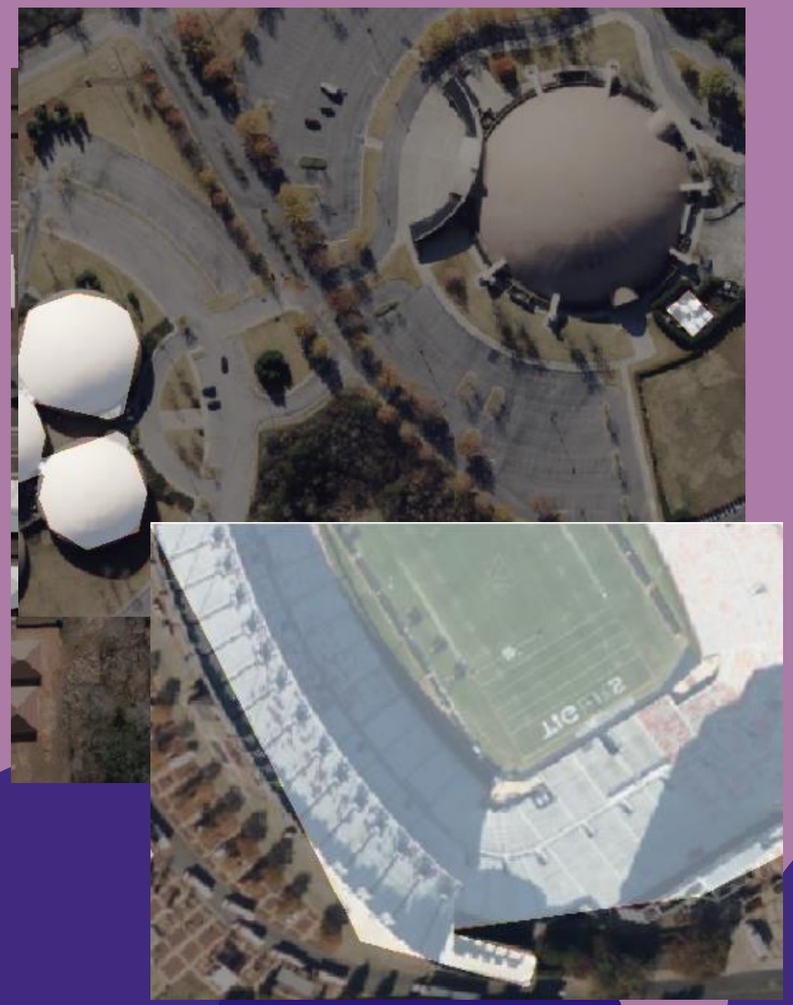
- 건물인데 마스킹이 되어 있지않거나 건물이 없는데 마스킹이 되어 있는 이미지
- 밀집된 건물에서 오류가 많았음
- 잘못된 마스킹의 갯수가 많은 이미지의 경우 훈련에 제외

일관되지 않은 마스킹 제거

- 같은 형태를 가지고 있지만 마스킹 되지 않은 건물
- special building ex)구형, 원기둥형 건물, 스타디움

PREPROCESSING

학습데이터 취사선택



잘못된 마스킹 이미지 제거

- 건물인데 마스킹이 되어 있지않거나 건물이 없는데 마스킹이 되어 있는 이미지
- 밀집된 건물에서 오류가 많았음
- 잘못된 마스킹의 갯수가 많은 이미지의 경우 훈련에 제외

일관되지 않은 마스킹 제거

- 같은 형태를 가지고 있지만 마스킹 되지 않은 건물
- special building ex)구형, 원기둥형 건물, 스타디움

PREPROCESSING

학습데이터 취사선택

잘못된 마스킹 이미지 제거

- 건물이 없는데 마스킹 된 경우 사람이 직접 보더라도 구분이 힘든 경우가 있었음

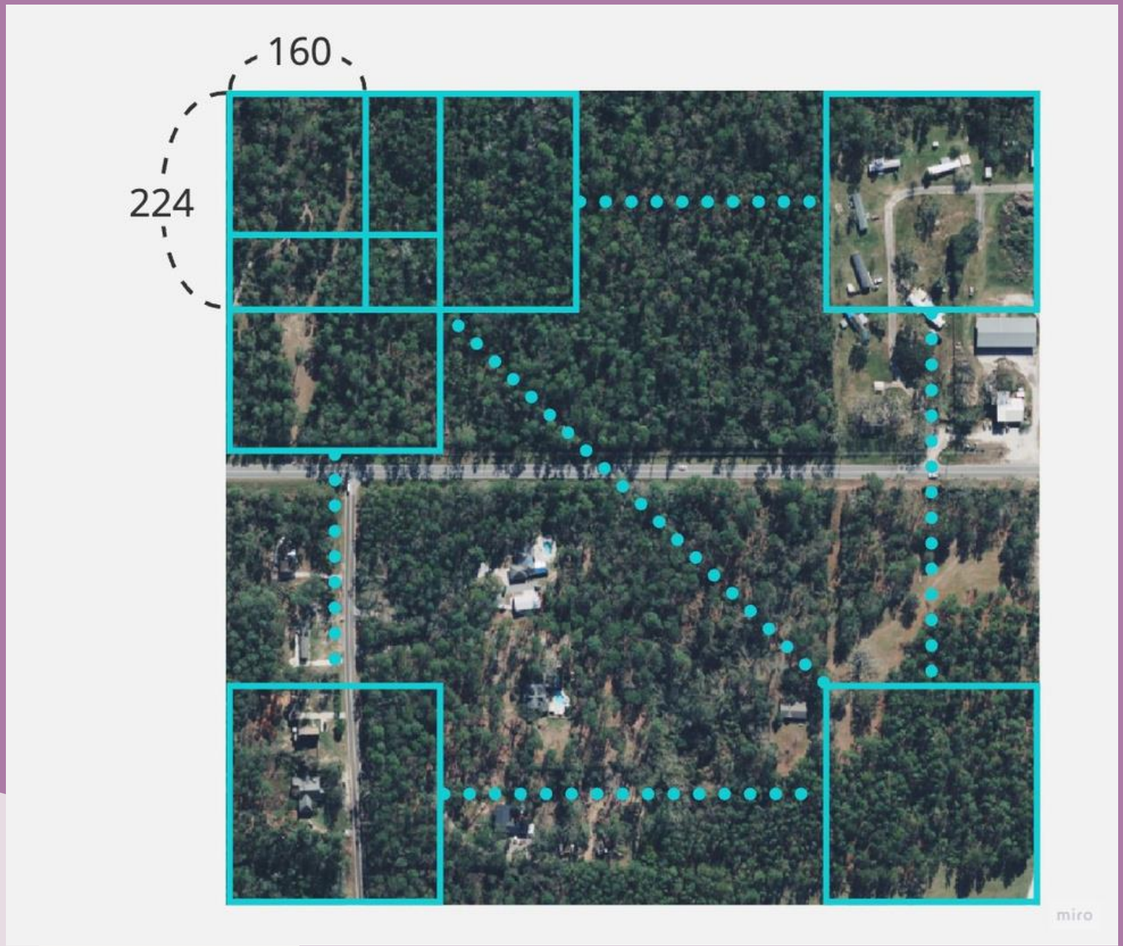
- 밀집된 숲 속 건물,
연속적으로 나열되어있는 건물의 경우
자세히 보지 않으면
마스킹이 제대로 되었는지 구분하기 힘듦

- 사진속 오류가 하나 존재 하더라도 정도가 심하지 않으면 학습 데이터로 사용함

- 직접 취사선택하는 단계에서의 한계점 이라고 볼 수 있음

PREPROCESSING

학습데이터 슬라이싱



224사이즈의 크기로 SIZE CROP

- Test사이즈와 같은 이미지로 사진 CROP
- Test이미지와 같이 잘린 건물, 건물이 없는 숲이나 도로 등과 같은 특징을 뽑아내기 적절함
- Test이미지와 해상도를 맞춤

160간격으로 사진 슬라이스

- 원본의 1024 사이즈를 160 간격의 간격으로 슬라이스 하여 숲, 도로, 건물의 위치에 대한 다양한 데이터를 얻을 수 있도록 하였음
- 한장당 36장의 224이미지를 생성하여 총 25만장 생성

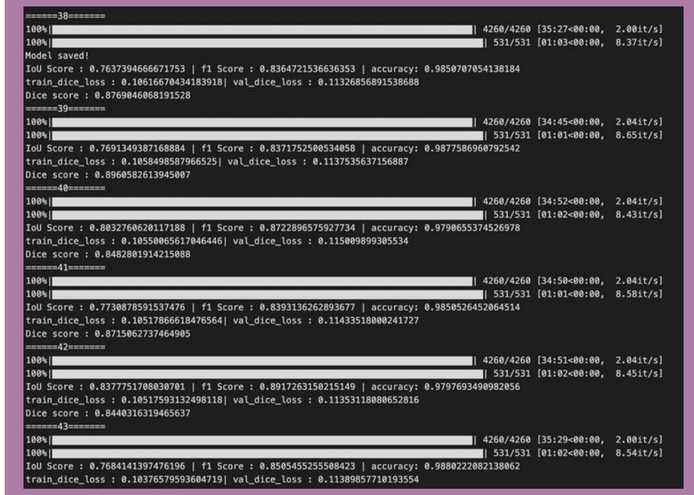
새로 TRAIN이미지 생성과 함께 CSV 생성

- 슬라이스한 이미지와 csv를 새로 생성하여 훈련에 사용

PREPROCESSING

VALIDATION SET

Step	Training Loss	Validation Loss
12000	0.036200	0.040822
15000	0.037000	0.040068
18000	0.035000	0.038748
21000	0.036800	0.038025
24000	0.036200	0.038506
27000	0.037000	0.037737
30000	0.047500	0.037172
33000	0.039100	0.037538
36000	0.028300	0.037430
39000	0.033300	0.037633
42000	0.033200	0.037128
45000	0.022000	0.037531
48000	0.026000	0.037217
51000	0.025200	0.036917
54000	0.024300	0.037030



슬라이싱된 데이터의 10%, 20%

- 슬라이싱된 25만장에서 10%, 20% 의 데이터 선택

OVERFITING 방지

- 학습 데이터에 과적합 되지 않도록 Validation Loss를 모니터링

STEP 2

데이터 증강

Augmentation

OVERFITTING

- 제한된 data set, image로 인해서 train data가 부족함
- train data에 대해 과하게 학습하는 overfitting이 발생함

AUGMENTATION의 종류

- Geometric Transformation
  - 가장 간단하고, 비용이 적은 transformation 중 하나
  - 데이터의 다양성을 늘리고, 건물의 불변 방향을 학습하는데 도움을 줌

- Pixel Transformation
  - 데이터의 다양성을 증대 시킴
  - 다만, 비용이 크고, 너무 많은 종류를 적용하거나 변형을 크게 하다 보면, 원본 이미지와 멀어져 오히려 성능이 떨어지는 부분이 있었음
  - OneOf를 사용해 한 가지만 적용하고, 확률 값을 배정함

Geometric Transformation

HORIZONTALFLIP, VERTICALFLIP

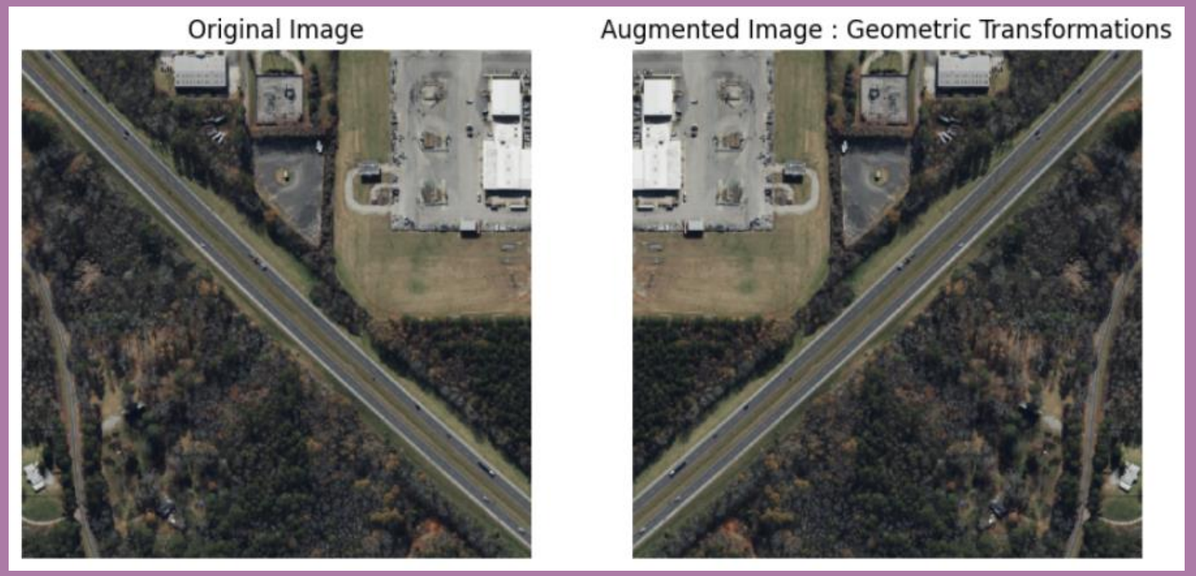
- 원본 이미지를 수평, 수직으로 뒤집음

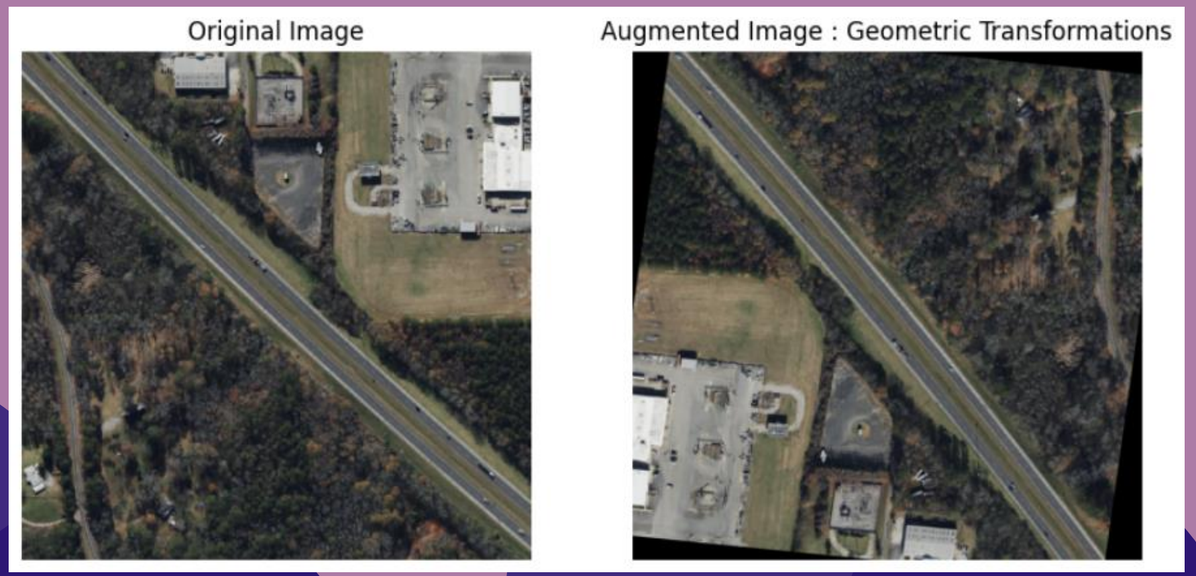
RANDOMROTATE90

- 원본 이미지를 0, 90, 180, 270 중 랜덤으로 회전시킴
- 건물의 불변 방향을 학습하는데 도움을 줌

SHEAR를 사용한 AFFINE 변환

- Affine 변환을 적용해서 원본 이미지를 x, y 방향으로 찌그러트림





Geometric Transformation

예측한 마스크 병합시 발견한 오류

- 같은 건물이라도 224 * 224 크기로 잘랐을때 상단에 위치 한 상황과 하단에 위치한 상황이 있다
- 건물이 상단에 위치해있을때는 건물을 제대로 인식 하지 못함
- 뒤집기, 돌리기 와 같은 증강 방법으로 이 문제를 해결하고자 함

삼원바이오비료영농조합법인-네이버지도

Pixel Transformation

CLAHE

- 이미지를 작은 블록으로 나누고, 히스토그램 평탄화를 적용함

RANDOM BRIGHTNESS CONTRAST

- 이미지의 밝기와 대비를 임의로 조정함

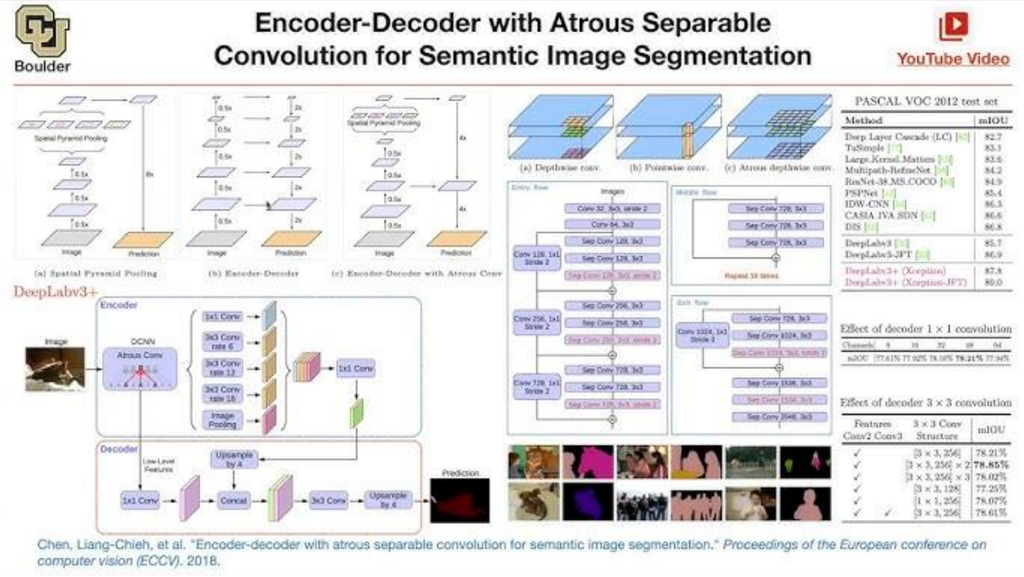
GAUSSNOISE

- 원본 데이터에 화이트 노이즈를 추가함

STEP 3

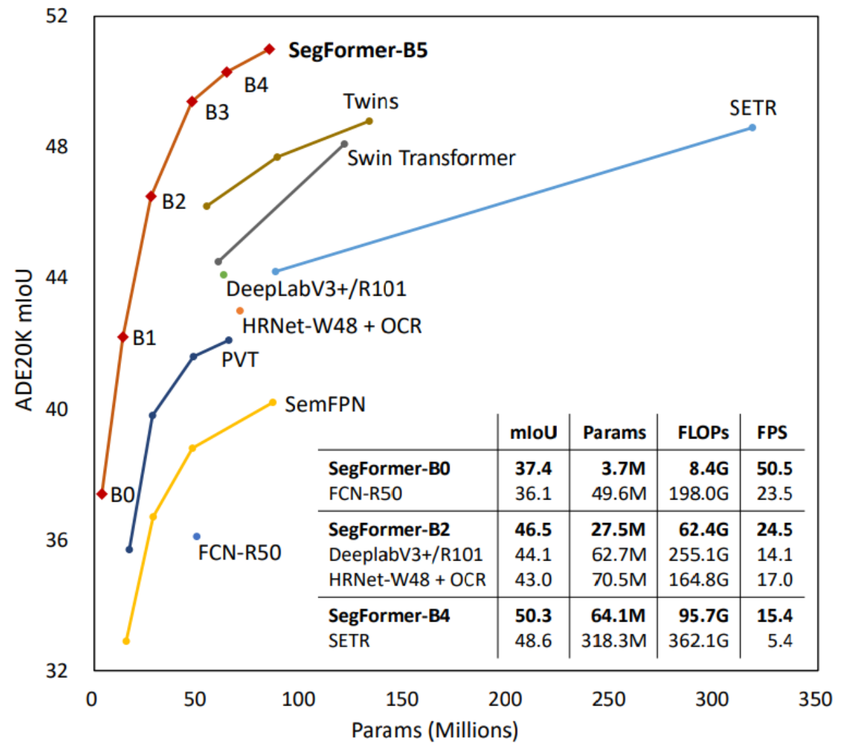
모델

DeeplabV3+ & exception71



EXCEPTION71 & DEEPLABV3+, DICE LOSS

- excpeiton71 인코더 & DeeplabV3+ 디코더를 사용
- exception71 인코더는 imagenet 가중치를 적용하여 사용
- diceloss를 적용하여 대회와 유사한 스코어, 손실을 적용함



	mIoU	Params	FLOPs	FPS
SegFormer-B0	**37.4**	**3.7M**	**8.4G**	**50.5**
FCN-R50	36.1	49.6M	198.0G	23.5
SegFormer-B2	**46.5**	**27.5M**	**62.4G**	**24.5**
DeeplabV3+/R101	44.1	62.7M	255.1G	14.1
HRNet-W48 + OCR	43.0	70.5M	164.8G	17.0
SegFormer-B4	**50.3**	**64.1M**	**95.7G**	**15.4**
SETR	48.6	318.3M	362.1G	5.4

Figure 1: **Performance *vs.* model efficiency on ADE20K.** All results are reported with single model and single-scale inference. SegFormer achieves a new state-of-the-art 51.0% mIoU while being significantly more efficient than previous methods.

segformer

트랜스포머 기반 모델

imagenet1k로 pretrain 된 모델(mit-b*) finetuning

최종 segmentation mask가 가로 1/4 세로 1/4로 줄어듦

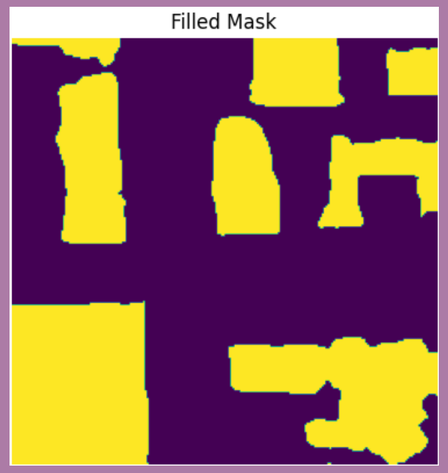
56 x 56 에서 224 x 224 크기로 늘려준다

b5 모델이 가장 성능이 좋았음

ade20k, citiscapes 데이터셋으로 finetuning된 모델 사용

STEP 4

후처리



Findcontour

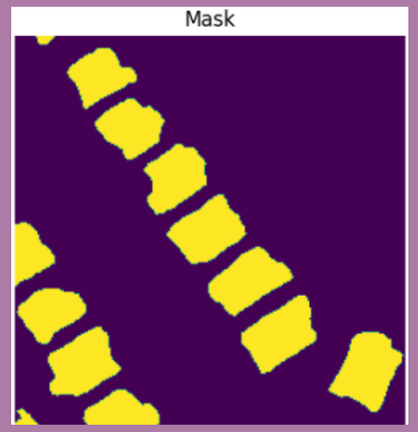
빈 공간 채우기 및 작은 객체 제거하기

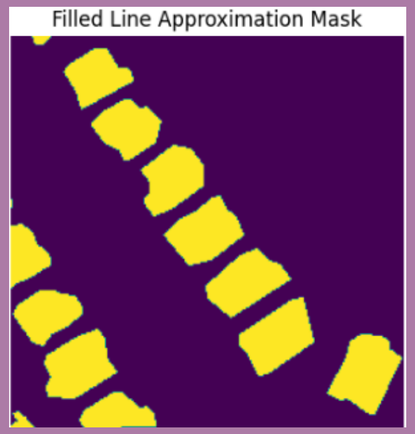
open CV에서 contour를 검출하는 함수

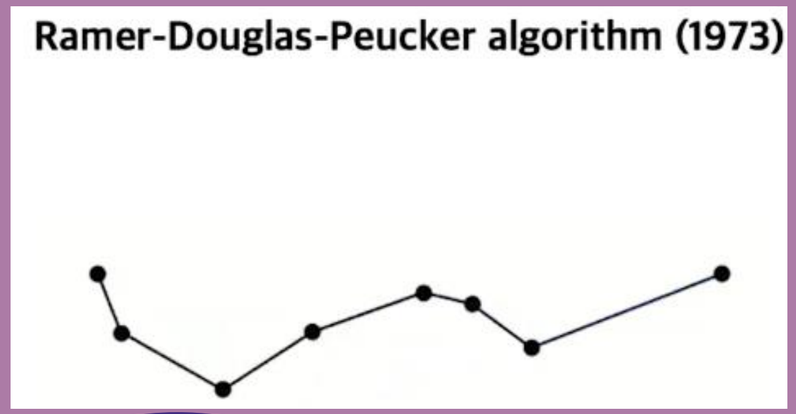
cv2.RETR_EXTERNAL를 이용하여 가장 바깥쪽 contour만 검출

contour 내부의 픽셀 수를 세서 특정 숫자 이하면 제거(작은 객체 제거하기)

마스크 경계를 그리고 내부를 채움(빈 공간 채우기)







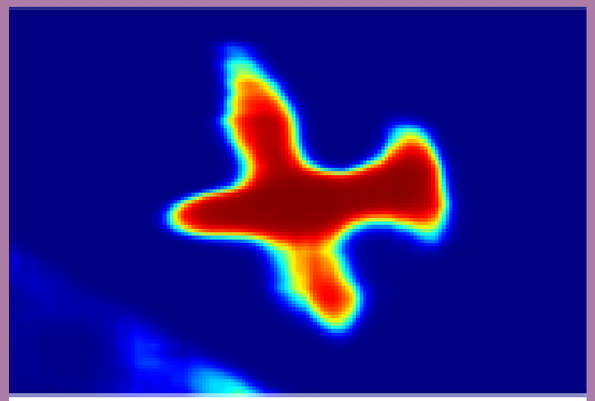
approxPolyDP

경계 직선화

학습 과정을 통해 얻어진 segmentation 마스크의 경계는 blur처리 된 것처럼 모호함

따라서 후 처리 단계에서 경계 직선화를 통해 실제 객체의 모서리나 선명한 특징을 되살릴 필요가 있음

maximum distance를 epsilon값을 기준으로 점을 지우는Douglas-Peucker알고리즘을 사용한 approxPolyDP함수를 사용하여 경계 직선화 구현



CNN output

CRF output

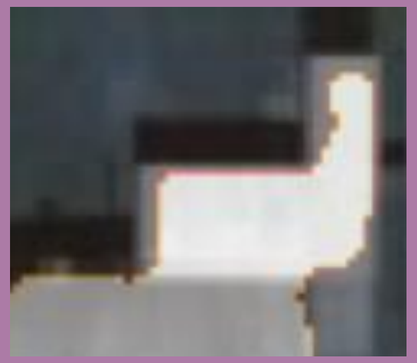
Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs-2016

모델 만으로 segmentation에서 정확한 윤곽을 예측하는 일은 어려움

논문에선 기존과 차별화된 fully connected CRF를 사용하여 복잡한 객체 경계를 정확히 포착 하도록 함

논문에서 사용된 모델이 DeepLab v2이고 우리가 사용한 모델이 DeepLab v3인 점을 미루어 봤을 때 결과가 개선될 가능성이 높다고 판단함

Before

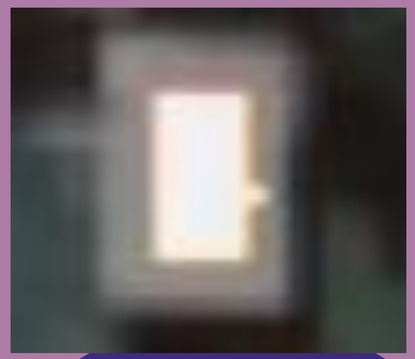


After

Morphology + CRF

노이즈 제거

- cv2의 erode, dilate 함수를 차례로 적용하여 노이즈를 제거함.
- dilate 함수를 erode 함수 뒤에 적용하여 객체를 본래 크기보다 팽창시킴으로써 가장자리의 빈 부분을 채움.



Before

After

객체를 본래의 크기에 맞게 정제

- dilate 함수를 통해서 본래의 크기보다 더 커진 객체가 존재함.
- CRF를 사용해서 주변의 RGB 컬러 비교를 통해서 본래의 크기에 맞게 조정함.

STEP 5

앙상블

ENSEMBLE

앙상블

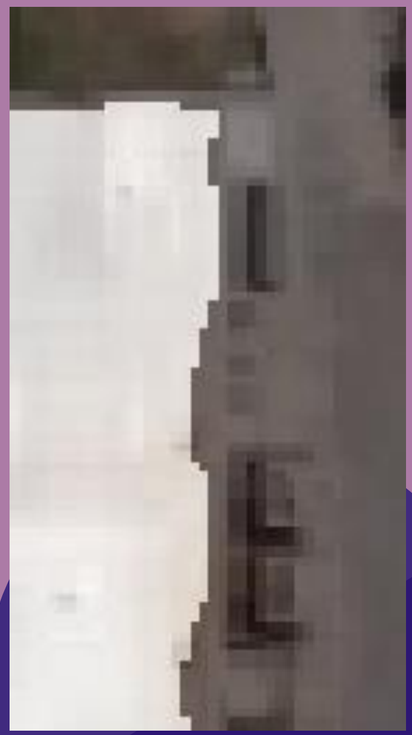
HARD VOTING

학습한 segformer, deeplabv3+ 모델의 결과를 가지고 하드보팅

방법이 간단하지만 성능향상을 가져옴





STEP 6

결과 분석

RESULT

마스킹 불균형

- 일반 주택의 경우 대부분 맞혔음.
- 큰 건물, 특이 건물에 대해서 잘 맞히지 못했음
- 같은 건물이라도 다른 클래스로 마스킹 했다면 더 좋은 결과를 볼 수 있을 것

느린 학습 속도

- 앙상블을 위해서 단일 모델이 아닌 여러 모델을 학습해야 함.

USAGE

현업에서의 사용

도시 개발, 인구 수 추정

- 우리 모델은 일반적인 가정 집에 대한 정확도가 높음.
- 건물들의 공간 분포 대한 정확한 정보 전달이 가능하므로 도시 계획 및 개발에 사용할 수 있음.
- 또한, 건물의 수량 또한 쉽게 알 수 있으므로 대략적인 인구 수 추정이 가능

자연 재해 등 재난 상황에서의 피해 파악

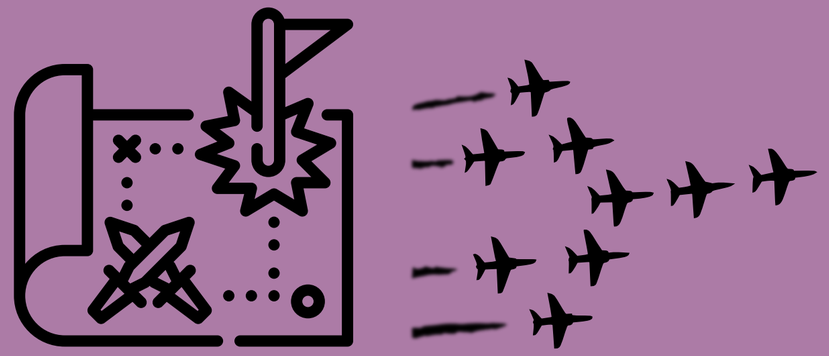
- 자연 재해 등 재난 상황에서 대략적인 피해의 정도를 파악할 수 있음.
- 같은 지점을 여러 시점에서 비교해서 피해를 입은 건물과 피해 정도를 대략적으로 판단할 수 있음.
- 인구 수에 대한 대략적인 예측도 가능하므로, 대략적인 인명 피해의 정도도 예측이 가능함.

USAGE

현업에서의 사용

군사적 전략 및 전술에서의 사용

- 적군에 대한 위치 정보를 얻을 수 있는 가장 쉬운 방법 중 하나가 위성 사진임.
- 적군의 건물에 대한 대략적인 위치 정보를 얻을 수 있음으로, 전투 전략에 도움을 줄 수 있음.
- 시가전이 중요해지는 현대전의 양상 속에서, building segmentation은 군사적 전략 전술 수립에 큰 영향을 미칠 수 있음.



개발 제한 구역에 대한 감시

- 건물 설치가 금지된 구역에 대한 건축물을 더욱 효과적으로 찾아낼 수 있음.
- 개발 제한 구역의 위성 사진에 건물로 마스킹 된 부분이 있는지 여부로 확인 가능.

USAGE

위성사진으로 지도 갱신

위성 사진으로 건물 인식

- 네이버, 카카오, 구글에서 제공하는 기본 지도는 위성 사진에서 새로 지어지거나 없어진 건물을 바로 반영하지 못함

- 우리 모델을 사용해서, 새롭게 지어지거나 없어진 건물을 인식할 수 있음

북 단양군 단성면 양당1길 75 일대 – 네이버지도

Thank you for your time.

동고동락