

거리뷰 AI스카우터 만들기

(거리뷰 영상에서 보석같은 데이터를 찾아내는 방법)



왕효선, 위동윤 NAVER / CLOVA
황호철, 한인규 NAVER / GLACE

CONTENTS

1. 도로 정보 스카우터란? (과제 정의)
2. 현실 세계의 데이터 분석
3. Annotation 데이터와 작업 효율화
4. Annotation 작업도구 만들기
5. 모델 학습과 평가
6. 스카우터 작동!
7. Q&A

1. 도로 정보 스카우터란? (과제 정의)

과제 정의

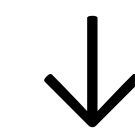


이미지 출처: 네이버 이미지 검색

26만 km의 도로

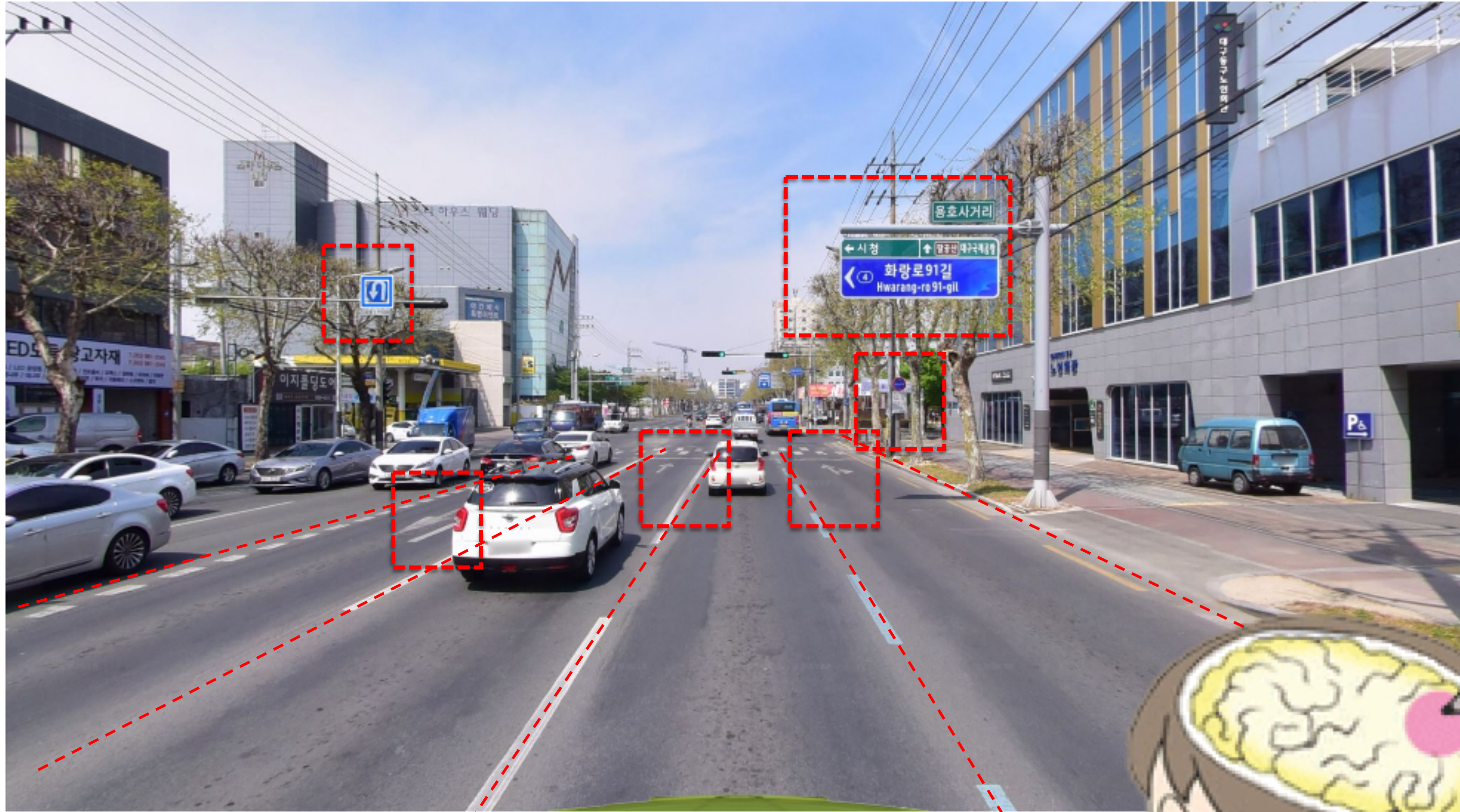
약 200종류의 표지판
수천 만개의 도로 안내 정보

업데이트 속도 = 유저의 만족도



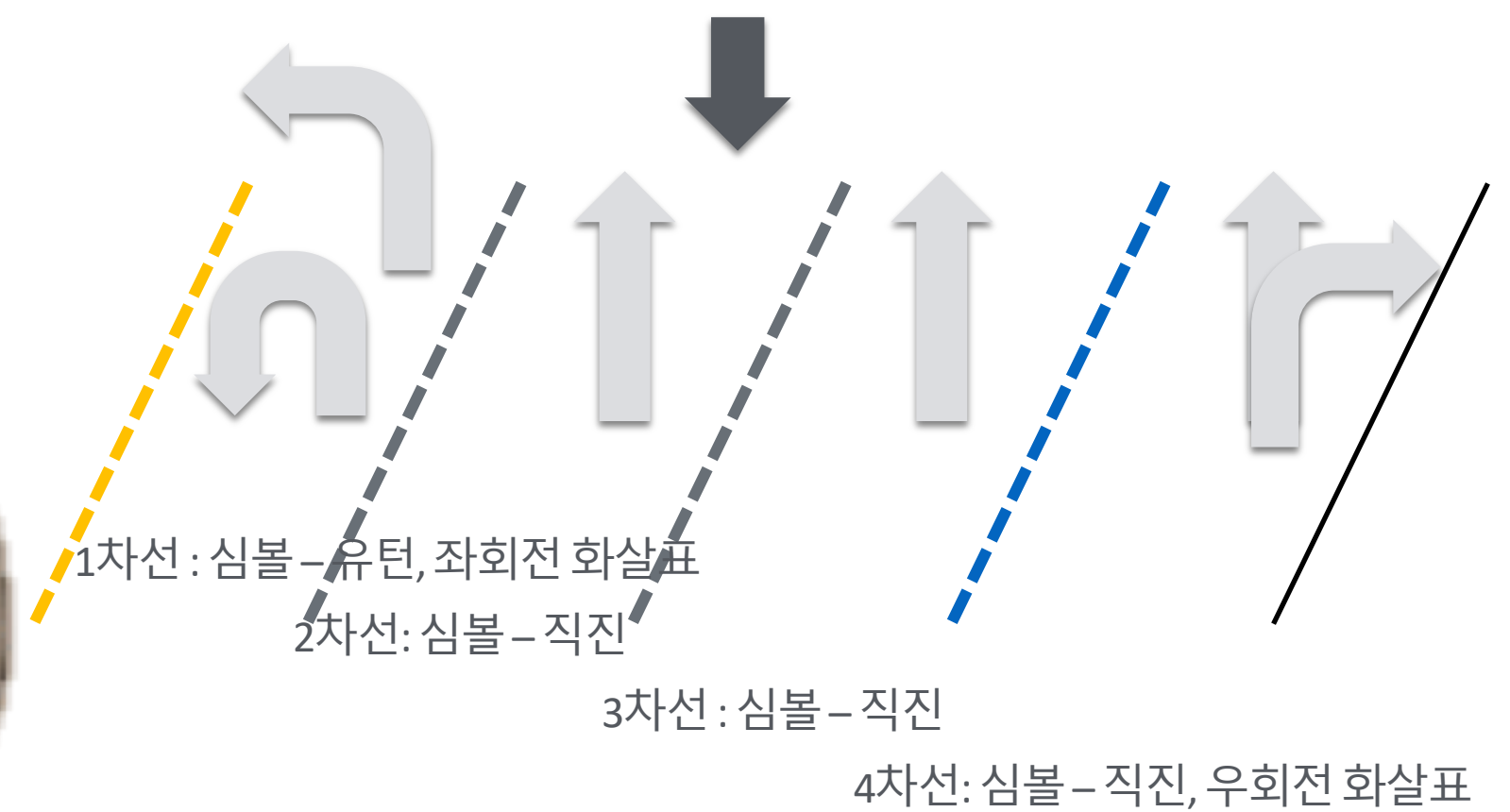
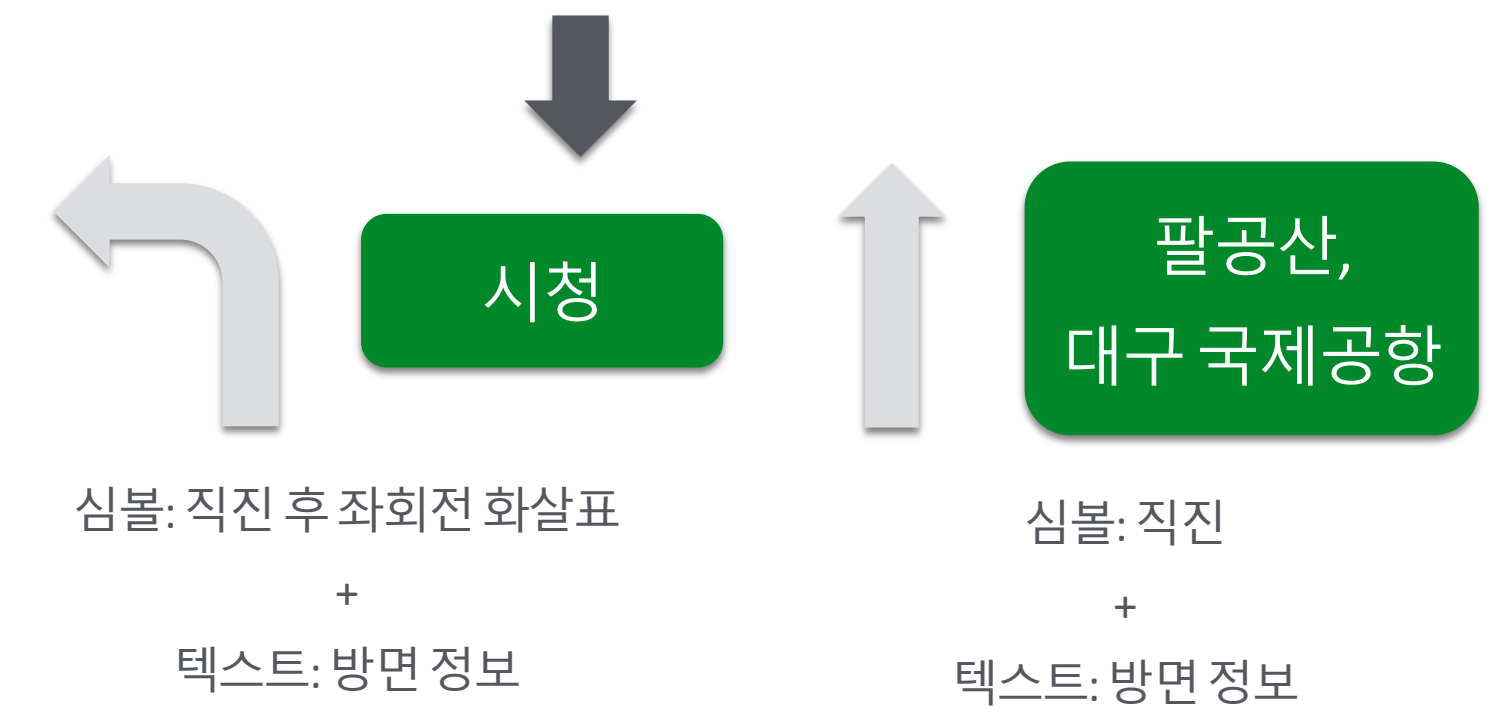
AI로 전국 구석구석의 도로 정보를
빠르게 인식하여 업데이트하는
스카우터를 만들수는 없을까?

과제 정의



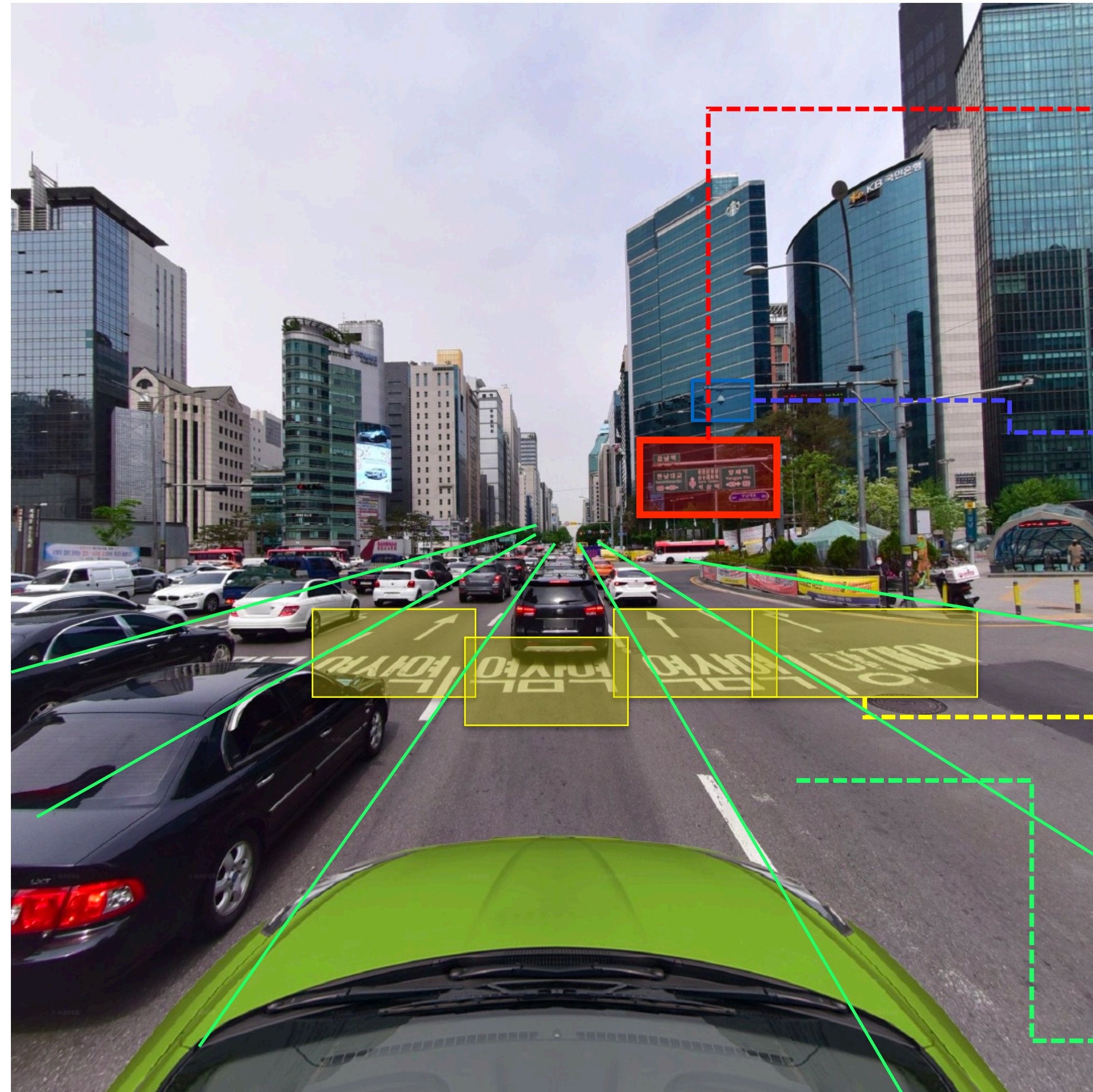
이미지 출처: 네이버 이미지 검색

- 방면 표지판이 있네
- 주의 표지판이 있네
- 보조 표지판이 있네



나는 시청 갈 거니까, 1차선으로 이동해야지

스카우터가 찾아내야하는 것



방면 표지판

분기점과 차선의 방면 정보를 통해 차량 운행 정보를 획득

안전 표지판과 시설물

과속 단속 카메라, 속도 제한, 키즈존 등 단속과 직결되는 주요 정보를 획득

노면 정보

방면 표지판과 함께 어떤 차선으로

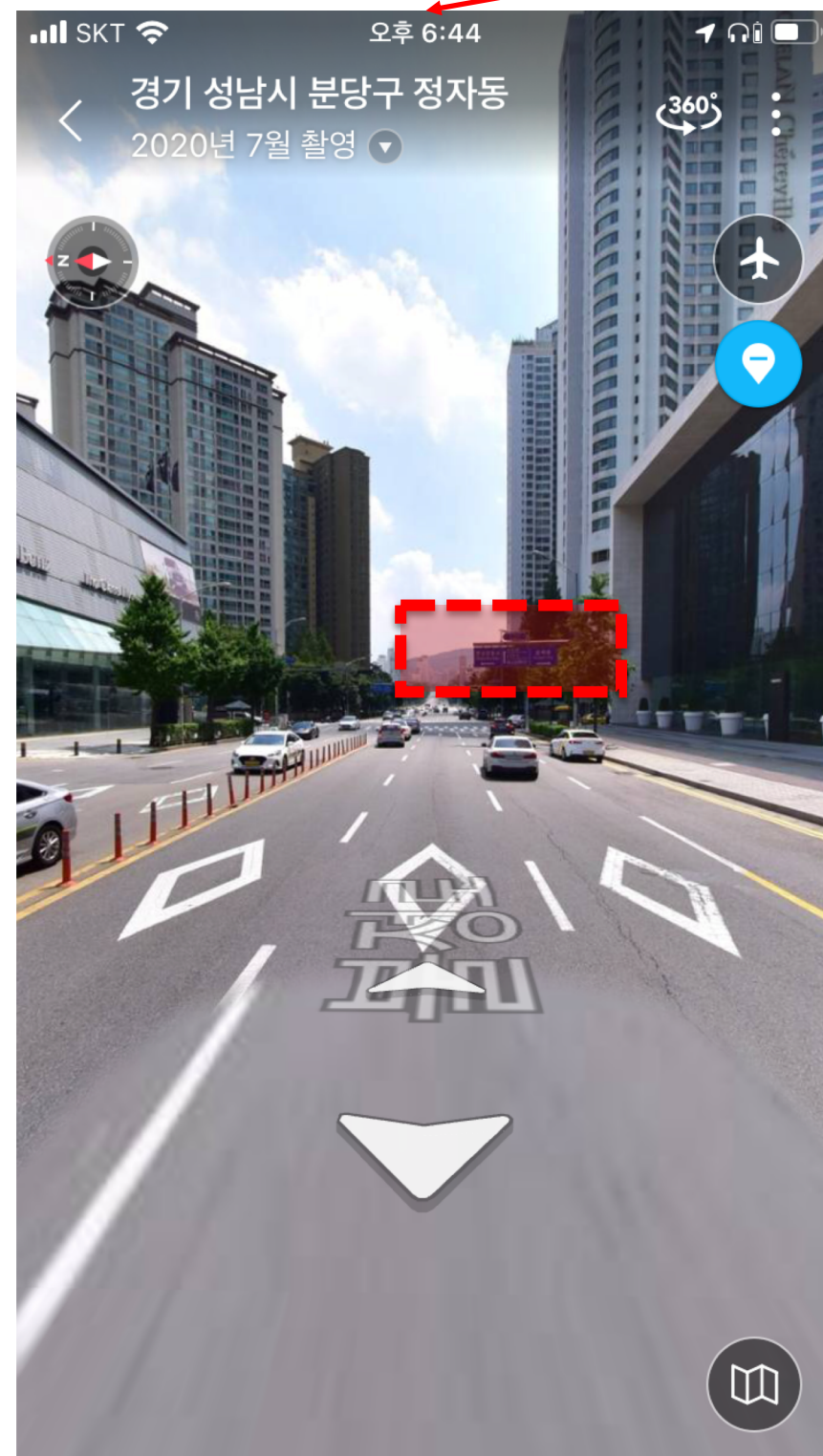
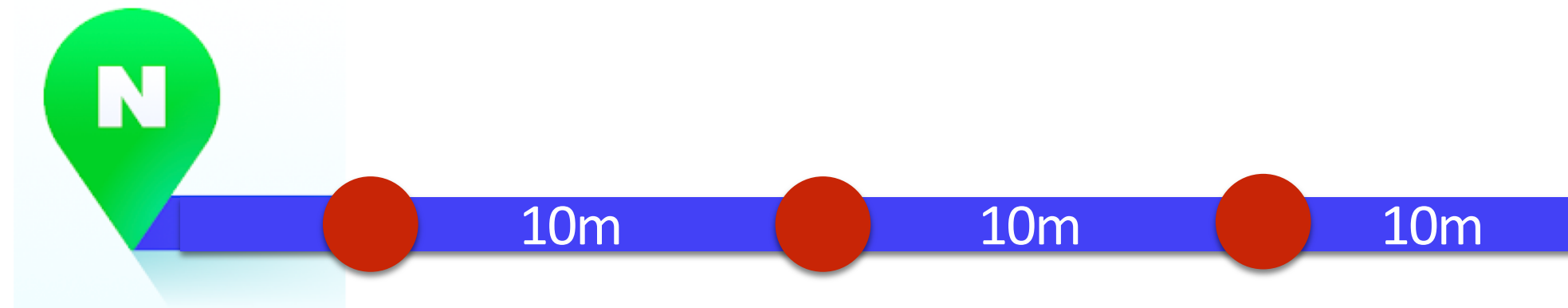
차선 정보

차선의 방면과 차선개수 등의 정보를 사용
교차로에서 방면과 다른 표지판 정보를 효과적으로 제외

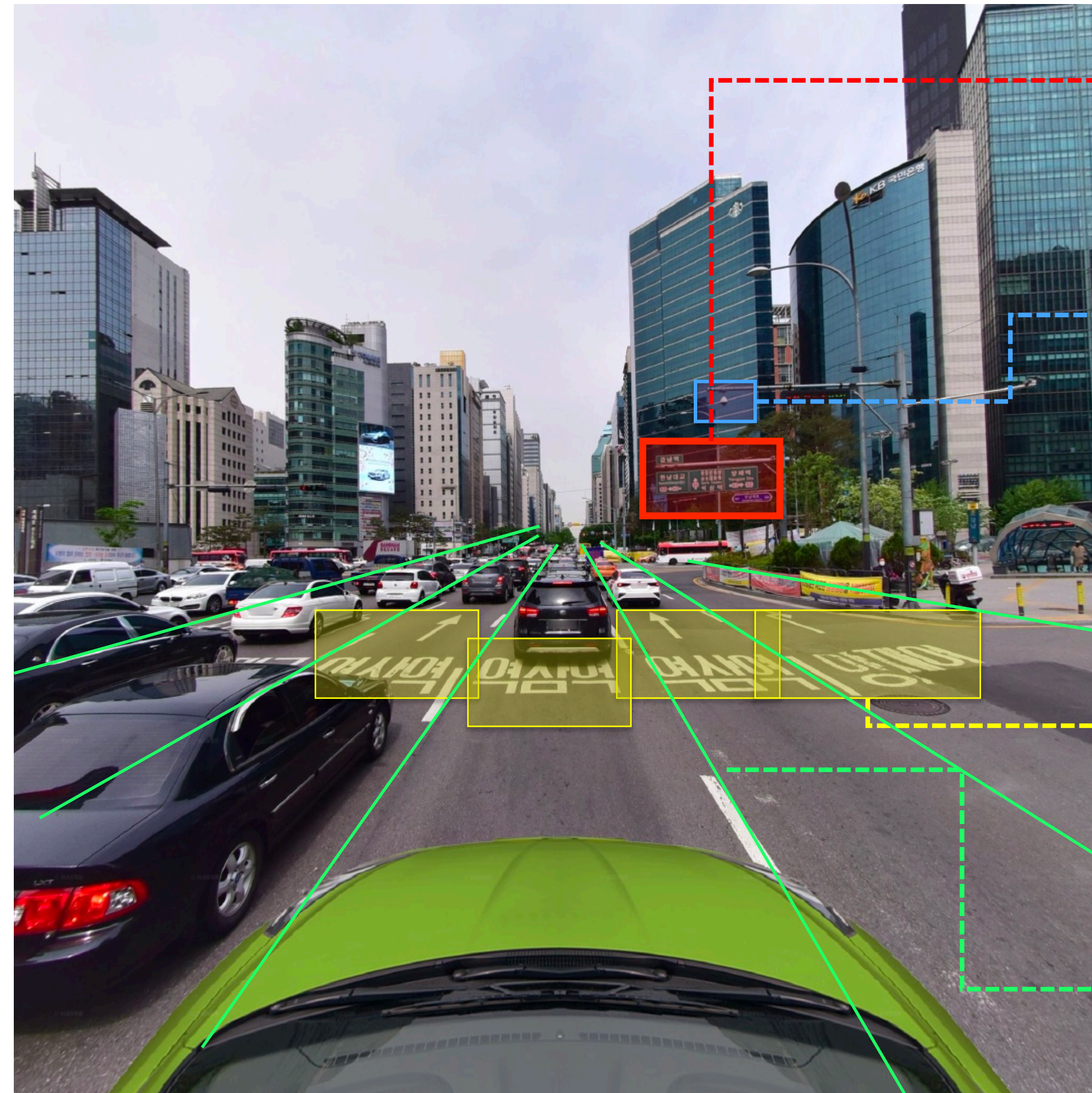
2. 현실 세계의 데이터 분석



우리에게 주어진 자료



우리에게 주어진 자료



1) 교차로의 직전-좌회전-우회전 3개 방향에 대한
방면을 안내하는 방면 표지판

2) 단속 카메라와 안내 표지판

3) 각 차선이 직진인지 회전인지를 안내하는 노면
정보

4) 중앙 차선 – 가장 바깥 차선까지의 6개 구분선을
통해서, 5개의 차선을 검출

1번 타자: 수많은 표지판과 정보 인식하기



이 표지판의 의미를 구분하실 수 있으신가요?

1번 타자: 수많은 표지판과 정보 인식하기



글씨와 심볼을 결합해야 정확한 의미를 해석 가능

1번 타자: 표지판 정보 인식하기

주의 표지		101 +자형교차로	102 T자형교차로	103 Y자형교차로	104 +자형교차로	105 +자형교차로	106 우회도로	107 우회도로	108 좌회도로	109 회전교차로	110 철길전널	111 우회교차로	112 좌회교차로	113 우회교차로	114 좌회교차로	115 2방향통행	116 오로막경사
117 내리막경사	118 도로폭이좁아짐	119 우측차로없어짐	120 좌측차로없어짐	121 우측방향통행	122 양측방향통행	123 중앙분리대시작	124 중앙분리대끝남	125 신호기	126 미끄러운도로	127 강변도로	128 노면고르지못함	129 과속방지턱	130 낙석도로	132 횡단보도	133 어린이보호	134 자전거	135 도로공사중
136 비행기	137 횡풍	138 터널	138의2 교량	139 야생동물보호	140 위험	141 상습정체구간	규제 표지		201 통행금지	202 자동차통행금지	203 화물자동차통행금지	204 승합자동차통행금지	205 이륜자동차및원동기 정지자전거통행금지	206 자동차·이륜자동차및 원동기정지자전거통행금지	207 경운기·트랙터및 손수레통행금지	210 자전거통행금지	211 진입금지
212 직진금지	213 우회전금지	214 좌회전금지	216 유턴금지	217 앞지르기금지	218 정차·주차금지	219 주차금지	220 차중량제한	221 차높이제한	222 차폭제한	223 차간거리확보	224 최고속도제한	225 최저속도제한	226 서행	227 일시정지	228 양보	230 보행자보행금지	231 위험물적재차량 통행금지
지시 표지		301 자동차전용도로	302 자전거전용도로	303 자전거 및 보행자 겸용도로	304 회전교차로	305 직진	306 우회전	307 좌회전	308 직진 및 우회전	309 직진 및 좌회전	309의2 좌회전 및 유턴	310 좌우회전	311 유턴	312 양측방향통행	313 우측면통행	314 좌측면통행	315 진행방향별통행구분
316 우회로	317 자전거 및 보행자 통행구분	318 자전거전용차로	319 주차장	320 자전거주차장	321 보행자전용도로	322 횡단보도	323 노인보호 (노인보호구역안)	324 어린이보호 (어린이보호구역안)	324의2 장애인보호 (장애인보호구역안)	325 자전거횡단	326 일방통행	327 일방통행	328 일방통행	329 비보호좌회전	330 버스전용차로	331 다인승차량전용차로	332 통행우선
333 자전거나란히 통행허용	보조 표지		401 거리	402 거리	403 구역	404 일자	405 시간	406 시간	407 신호등화 상태	408 전방우선도로	409 안전속도	410 기상상태	411 노면상태	412 교통규제	413 통행규제	414 차량한정	415 통행주의
415의2 충돌주의	416 표지설명	417 구간시작	418 구간내	419 구간끝	420 우방향	421 좌방향	422 전방	423 중량	424 노폭	425 거리	427 해제	428 견인지역	표지판류		주의 규제 지시 보조		
충돌주의	타설길이 250m	구간시작 200m	구간내 400m	구간끝 600m	→	←	↑ 전방 50M	3.5t	3.5m	100m	해제	견인지역	표지판류		주의 100-210	규제 100-210	지시 100-210

1번 타자: 표지판 정보 인식하는 “효과적인 모델”

표지판 정보 분해



분해 방법

표지판 종류 (5종)

X

내부 기호 (100종)

X

내부 텍스트

=

다양한 표지판의 의미

분해 효과

모델 재개발 없이
클래스를 추가하면

간단히 확장가능

X

새로운 기호 추가 시
기호만 따로 학습 가능

=

재사용성 개선

1번 타자: 표지판 정보 인식하는 “효과적인 모델”

현실 세계의 표지판



표지판 분해

일반 표지판

- 심볼: 카메라
- 텍스트: 단속중
- 텍스트: 250m

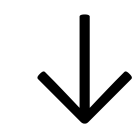
+

규제 표지판

- 텍스트 70

필요한 최종 결과

위치 상으로
가까운 표지판들끼리 조합



최고 속도 70km로
단속중인 카메라가
250m전방에 있다

→ 모델 1) object & symbol detection + OCR

1번 타자: 표지판 정보 인식하는 “효과적인 모델”

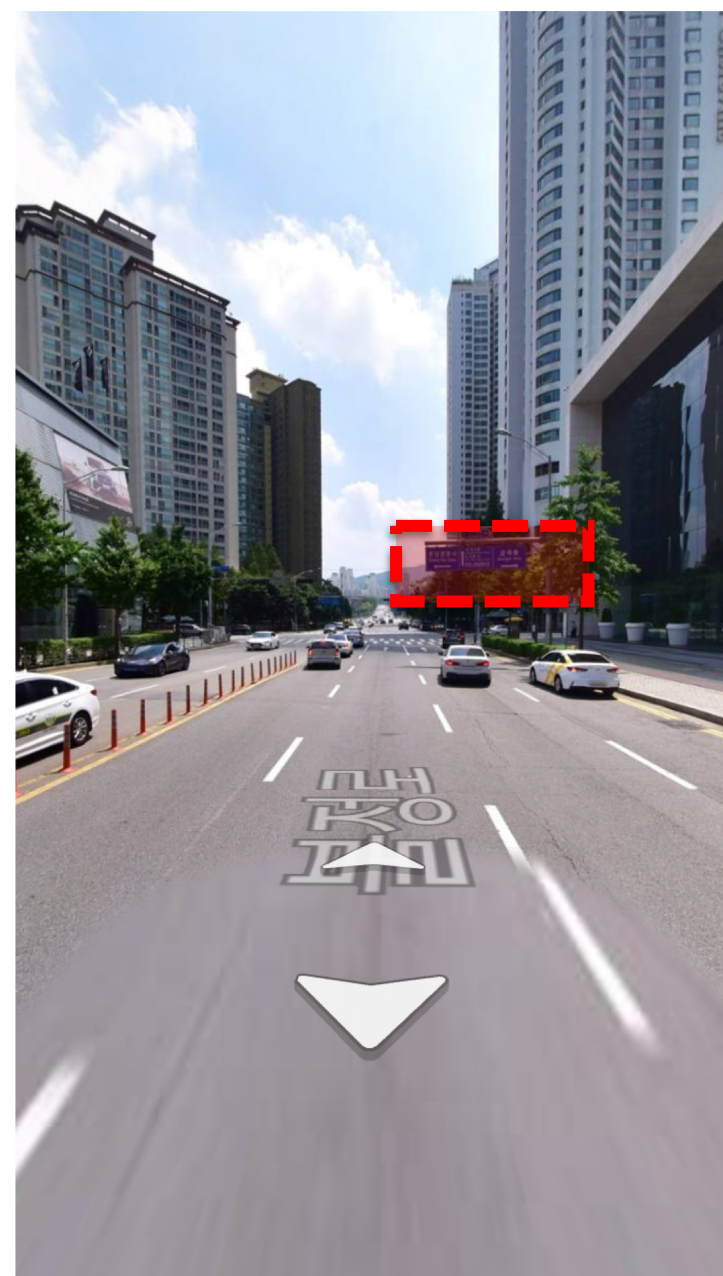
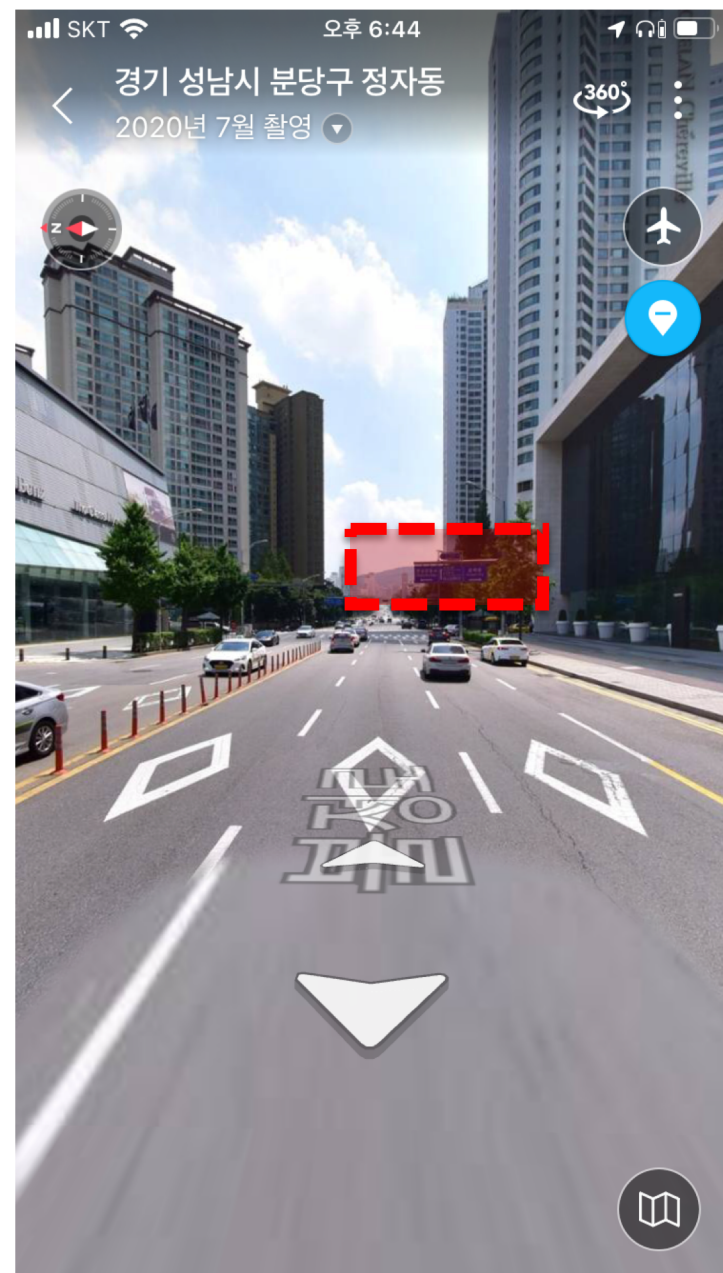


조금씩 다른 텍스트 x 심볼 x 웨입 x 조합이 대응가능한 모델

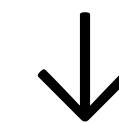
2번 타자: 동일한 객체임을 판단하기

거리뷰 속의 표지판

어떻게 위치를 인식할까



“표지판은 계속 등장합니다.
그렇다면 언제 단속 카메라가
있다는 것을 알려줘야 할까요?”



마지막 등장 위치로 추정

→ 모델 2) multi-object tracking

3번 타자: 차선 정보 활용하기

차선별 인식해야 하는 정보



차선별 인식해야 하는 정보

1. 차선 갯수
2. 진행 방향
3. 회전 정보
4. 방향 정보

3번 타자: 차선 정보 활용하기

차선별 인식해야 하는 정보



어떻게 차선 정보를 인식할까

1. 차로 구분

+

2. 차로 내 교통정보 인식

+

3. 방면 표지판 인식

=

방면 표지판 매칭

차선별 인식해야 하는 정보

- 교통 정보 인식
- 교통 정보 위치 추정
- 차선 인식
- 표지판 매칭

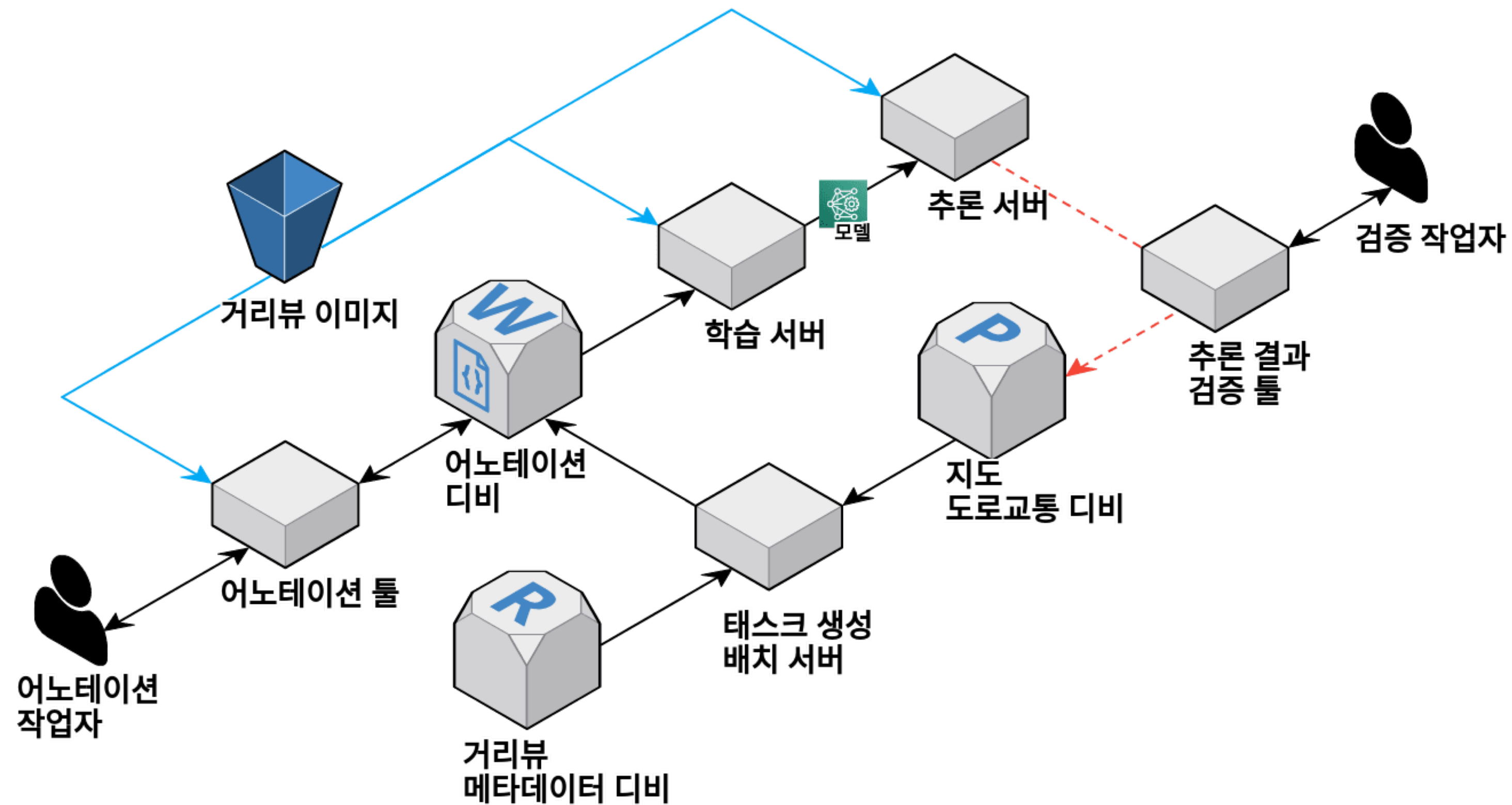
필요한 모델

1. Object& Symbol detection + OCR
2. Multi object tracking
3. Lane detection
4. Parsing

3. 어노테이션 작업

시스템 개요

- 거리뷰 이미지 특성에 맞는 전용 어노테이션 툴 개발
- 신뢰할 수 있는 추론 결과 반영을 위한 모델 검증 작업



어노테이터에게 할당할 태스크 만들기

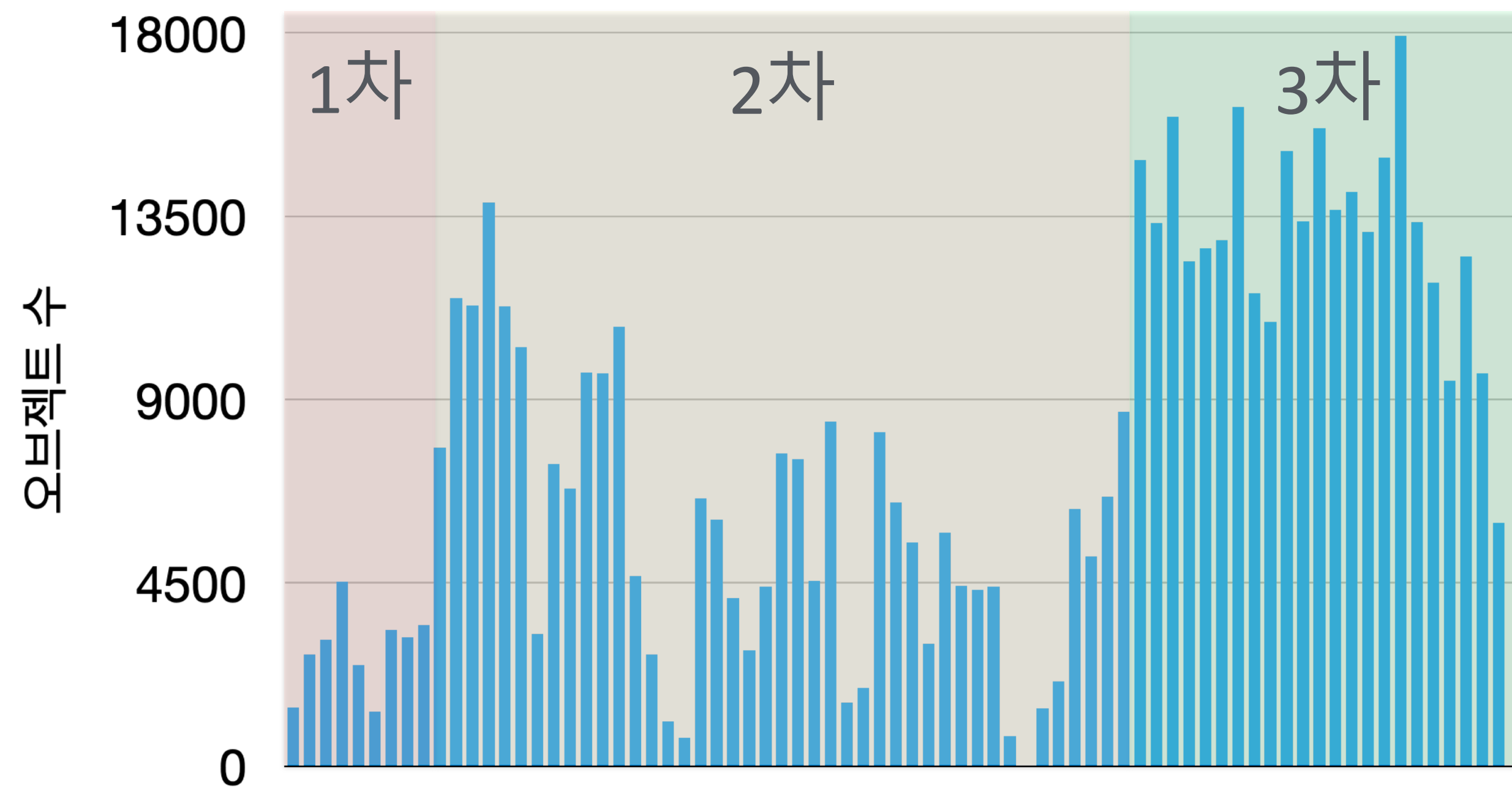


- 어노테이션 작업의 단위
- 도로 구간에 대한 연속적인 거리뷰 이미지
- 정보 많은 곳 주변
 - 교차로
 - 표지판
 - 카메라

어노테이터에게 할당할 태스크 만들기

효율적인 작업을 위한 노력

일별 작업량



1차: 거리뷰 이미지 100장씩

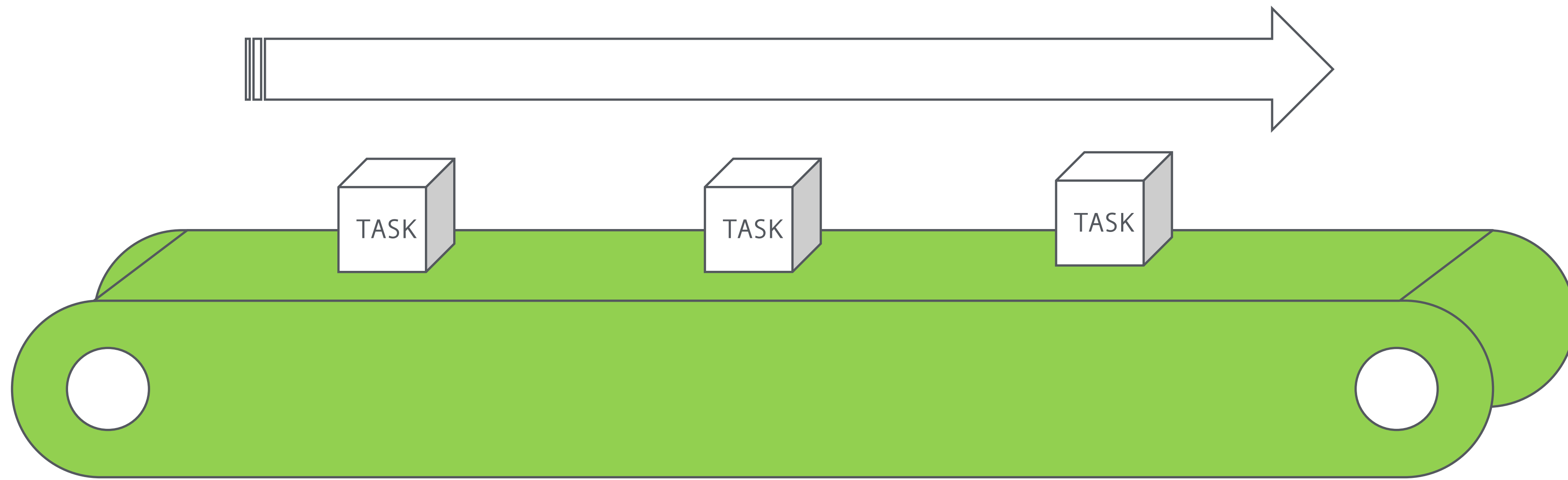
2차: 거리뷰 이미지 50장씩

3차: 추론 결과 반영 (Pseudo Labeling)

4. 어노테이션 툴

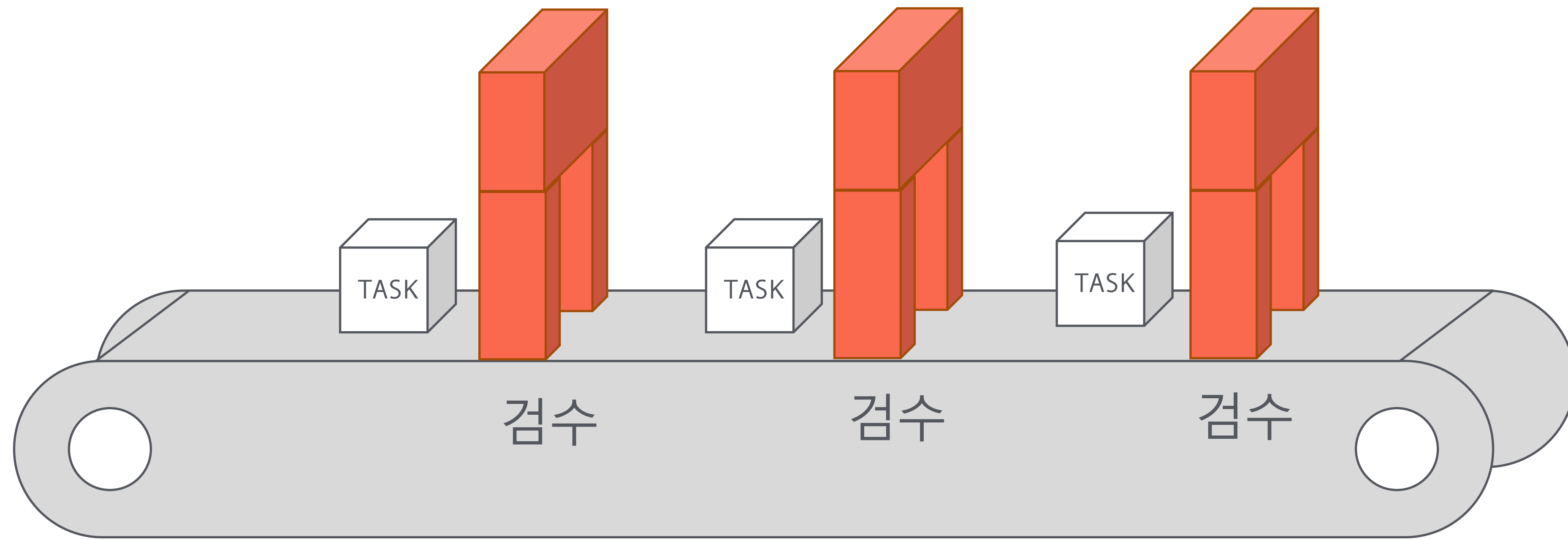
해결해야하는 과제

1. 복잡한 어노테이션 (표지판 분류, 차선분류, OCR등)
→ 단순 / 반복작업



해결해야하는 과제

2. 작업자의 숙련도가 모두 다름
- > 검수과정 강화

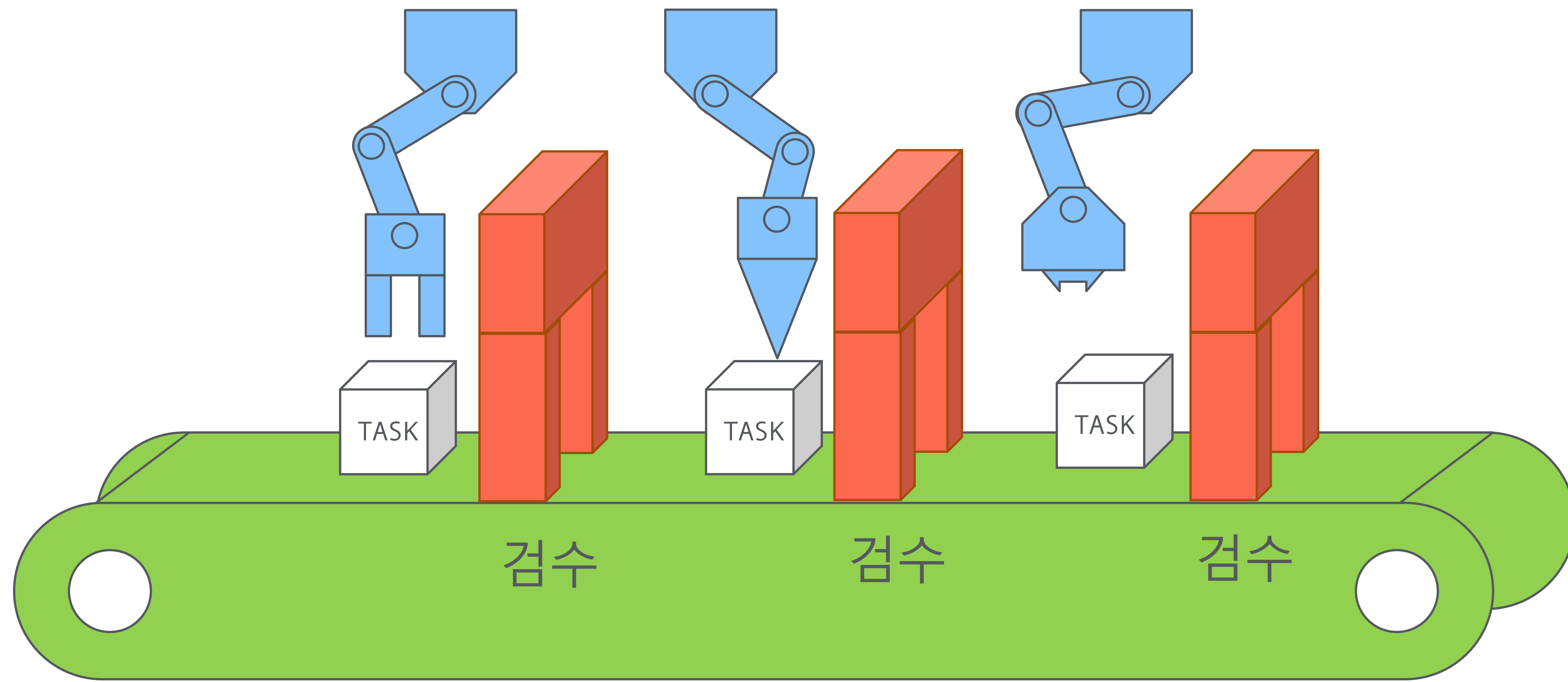


컨베이어 시스템

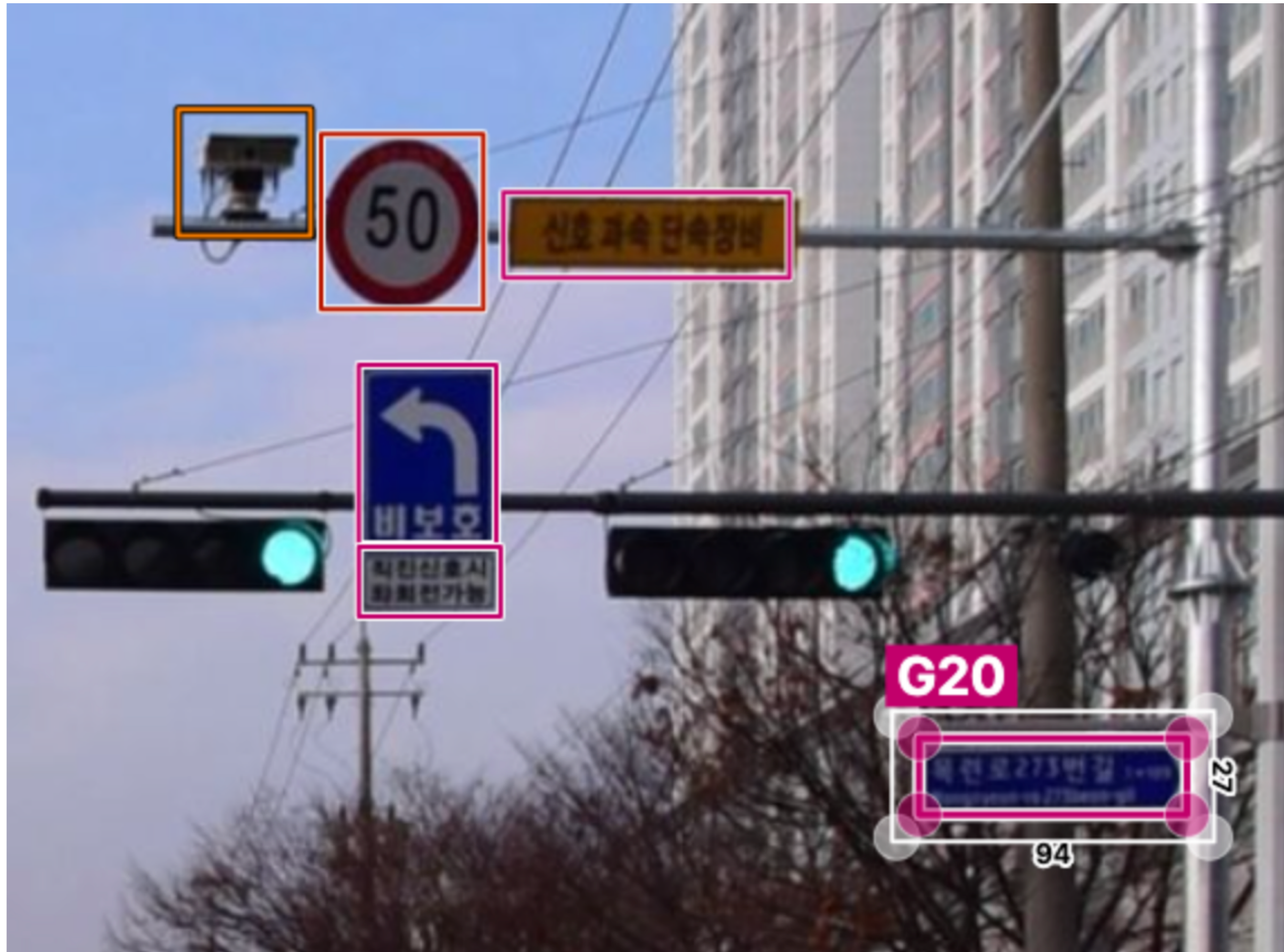
1. 찾기

2. 기술하기

3. 연결하기



Step1 찾기



일반표지 1

규제표지 2

주의표지 3

지시표지 4

노면정보 5

카메라 6

보조카메라 7

카메라박스 8

분할 9

기타 0

중앙선 Q

일반차선 W

외곽선 E

추정선 R

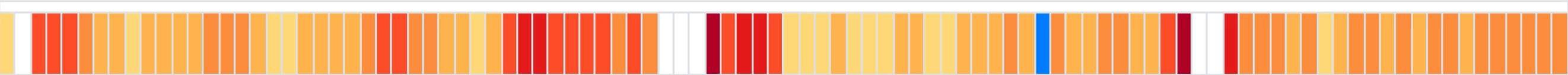
버스전용 T

기타차선 Y

332227 강변로



슬라이드 토클



선택된 오브젝트가 없습니다. Alt + 드래그 or 드래그 하여 화면이동, 마우스 휠 로 줌인/줌아웃

전체

표지판

노면

차선

카메라

분할/기타

모두열기 (느림)

모두닫기 (권장)

207x140			
	G61	속성변경	삭제
202x83			
	R194	속성변경	삭제
49x58			
	R195	속성변경	삭제
137x58			
	R93	속성변경	삭제
162x59			
	CL121	속성변경	삭제
1090x460			

332228

(10) 열기

332229

(7) 열기

332230

(9) 열기

332231

(10) 열기

Step2 기술하기

타입*

기타		
화살표	도로	물체
교통수단	도로종류	노면정보

교통수단*

 자전거	 이륜차	 자동차	 트럭	 버스
 기차	 비행기	 트랙터		

화살표*

 회전	 위	 아래	 왼쪽	 오른쪽
 유턴	 1시	 5시	 7시	 11시
 우좌 이중	 좌우 이중	 P턴	 X	

방향*

가로	세로
----	----

종류*

시간	명칭	거리	속도	중량	교통
----	----	----	----	----	----

내용*

횡단보도

디텍션에서 마킹한 오브젝트를
구체적으로 기술하는 작업

텍스트 1

심볼 2

ALL !

현재 44 / 완료 998 / 전체 1000

이전 <

다음 >

이전 ! ⬆

다음 ! ⬆

D10 !

속성변경

지시 표지판

☐ DC (인식가능한 오브젝트가 없음)

텍스트

심볼



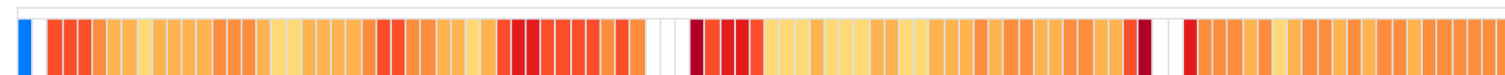
Step3 연결하기

연속되는 이미지에서 나타나는 오브젝트를 하나의 그룹으로 연결

332161 강변로



슬라이드 토클 A, D, ←, →



새 트래킹 시작 W

모든 오브젝트 보기 | 트래킹 된 오브젝트만 보기

CAM2

편집

전체삭제

CAM7

편집

전체삭제

C3

편집

전체삭제

C4

편집

전체삭제

C7

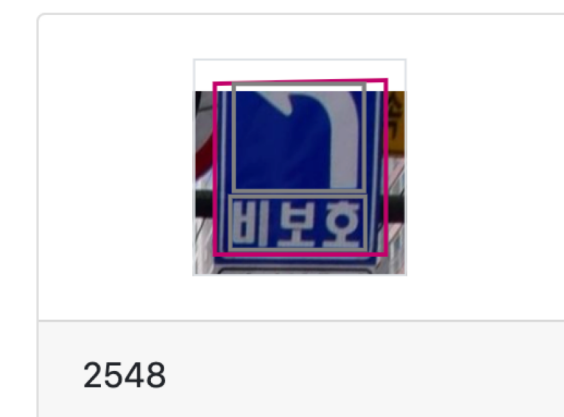
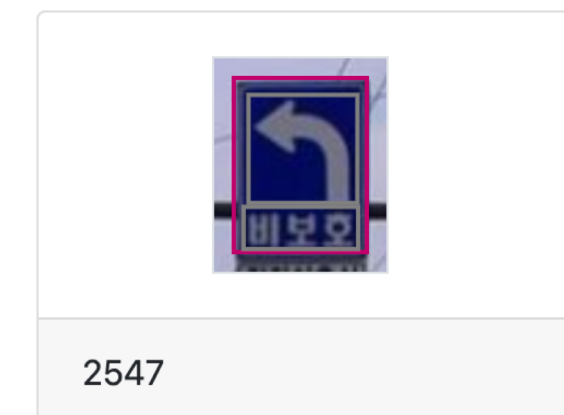
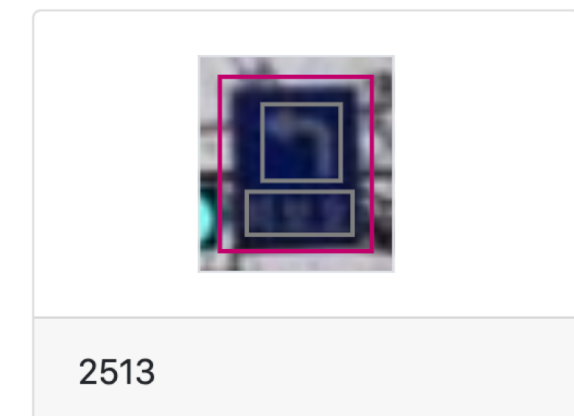
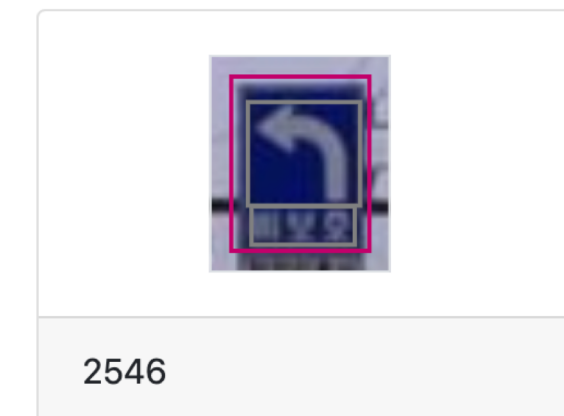
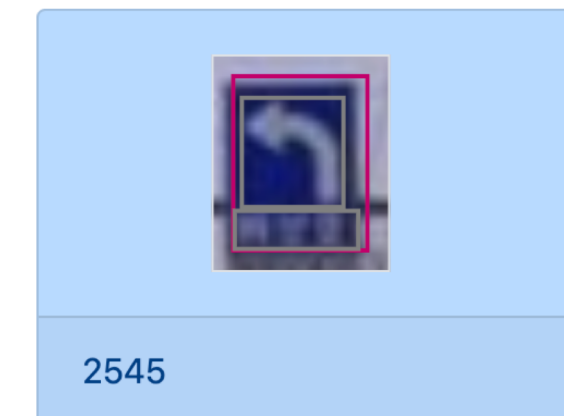
편집

전체삭제

C8

편집

전체삭제





슬라이드 토글 A,D,←,→



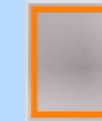
새 트래킹 시작 W

모든 오브젝트 보기

트래킹 된 오브젝트만 보기

CAM1

편집



CAM2

편집



CAM3

편집



CAM4

편집



CAM5

편집



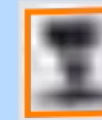
CAM6

편집



CAM7

편집

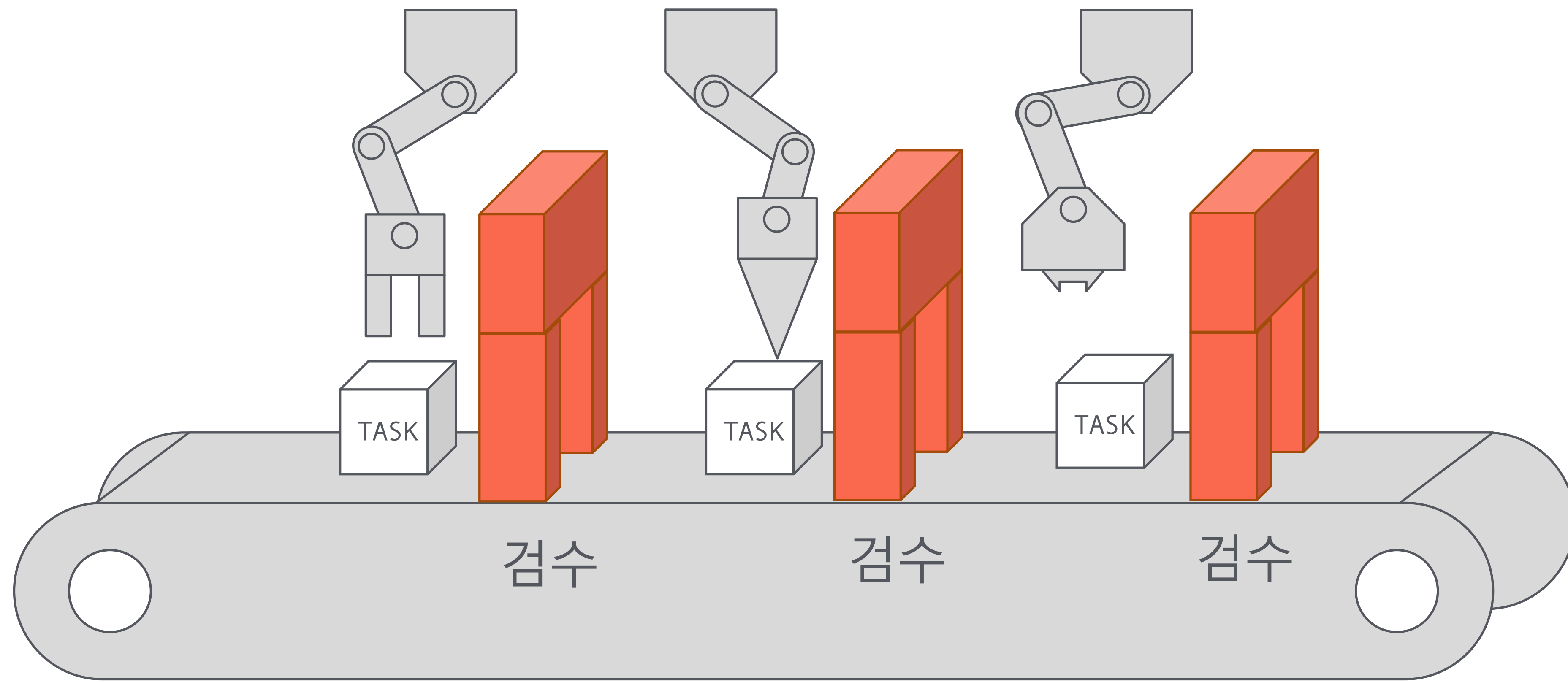


검수학기

1. 찾기

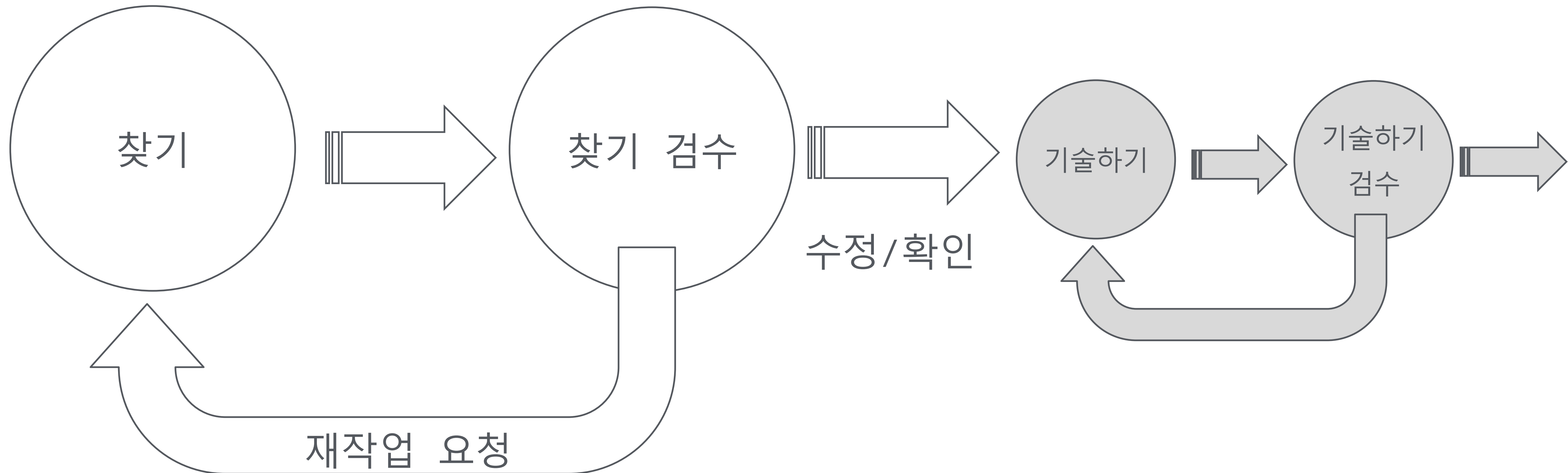
2. 기술하기

3. 연결하기



검수하기

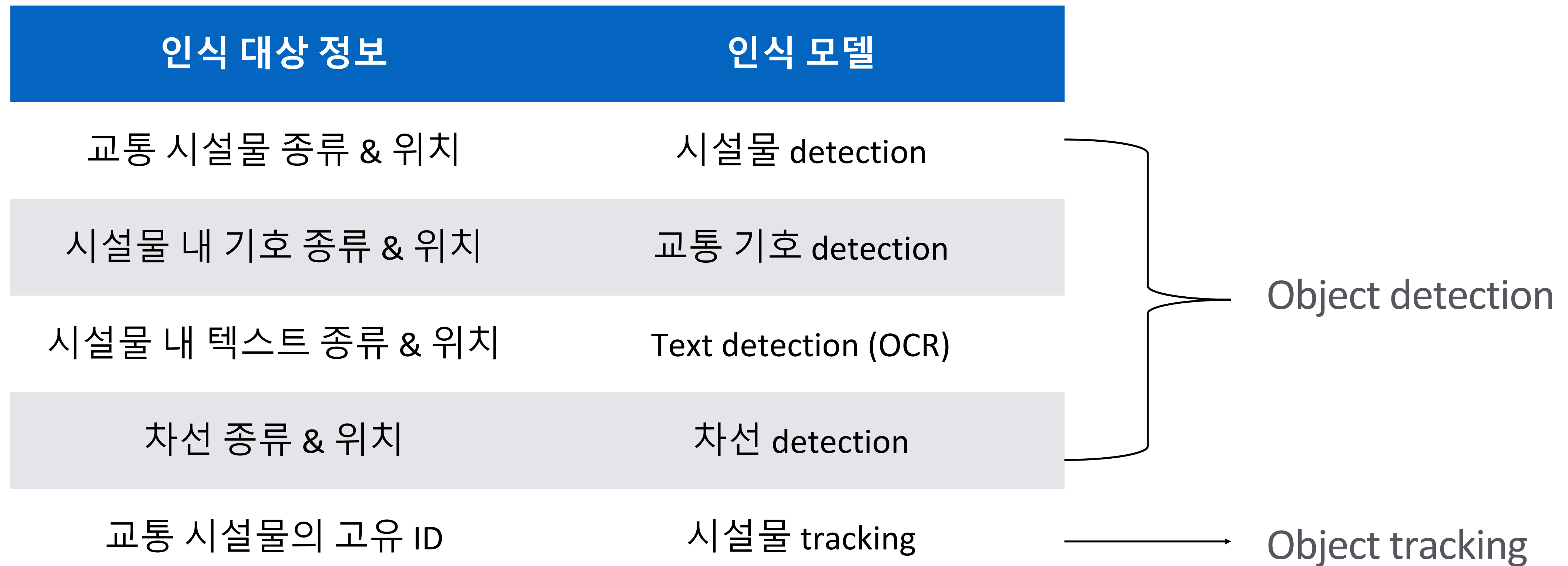
검수 과정에서 확인, 수정, 재작업 요청



5. 모델 설계

5.1. 인식 대상과 모델

- Object detection 모델 (4개) + Object tracking 모델 (1개)



5.2. Object detection

- 차선, 교통시설물, 노면정보, 시설물 내 기호 검출



5.2. Object detection

- 서비스/데이터 특성에 따른 모델 요구사항

서비스 특성

- 배치 서비스가 가능한 형태
- 유연한 서빙 환경 (CPU, GPU, 모바일..)

데이터 특성

- 인식 대상 별 크기 편차가 심함

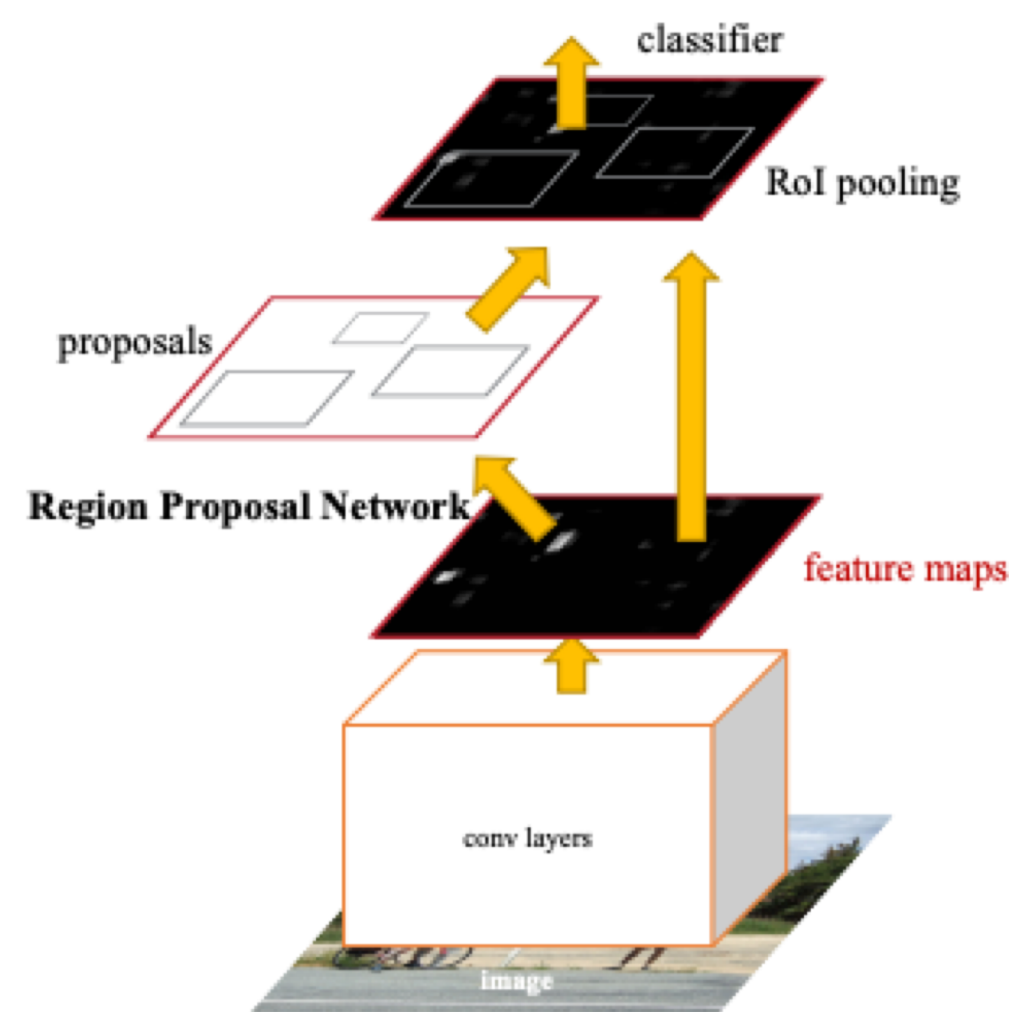


모델 요구사항

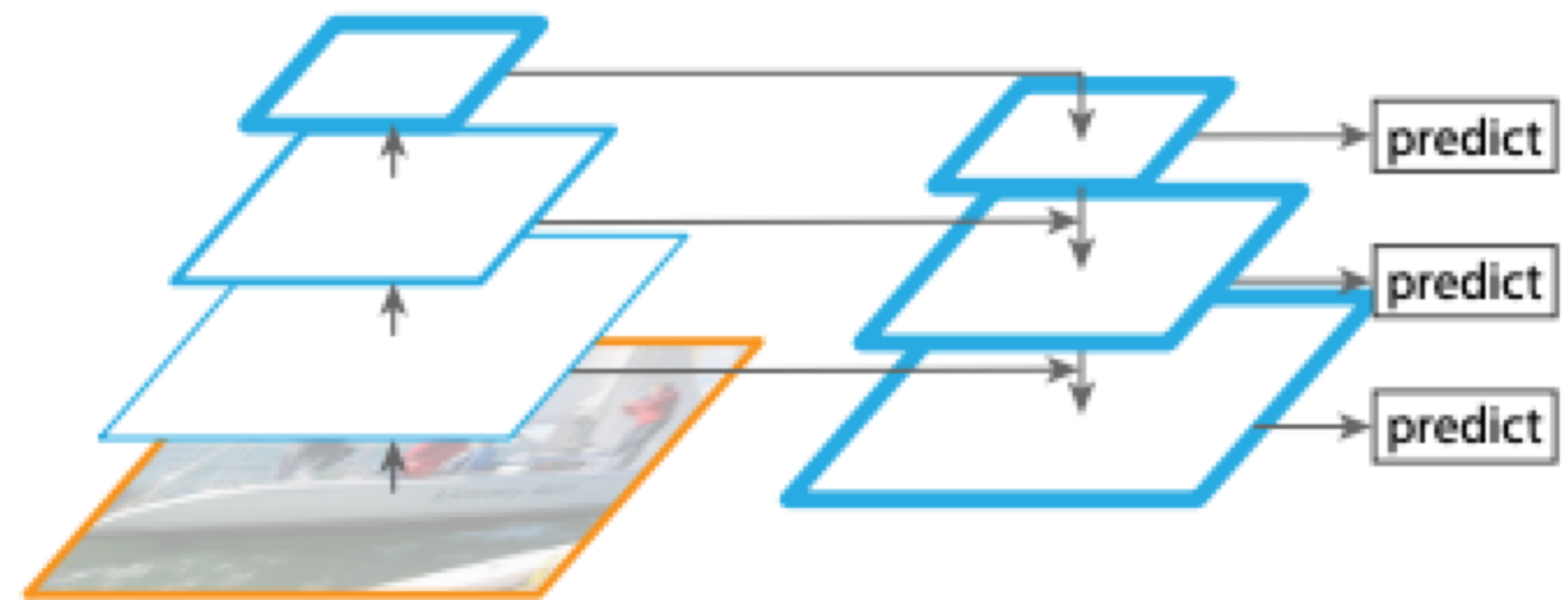
- 계산량 절감 < 정확도
- 다양한 크기의 객체 검출에 유리

5.2. Object detection – Baseline 모델

- Faster R-CNN with ResNet-FPN
 - Faster R-CNN: 정확도 향상을 위한 two-stage object detection
 - ResNet-FPN: 다양한 크기의 객체 검출을 위한 multi-scale feature



Faster R-CNN 구조 [1]



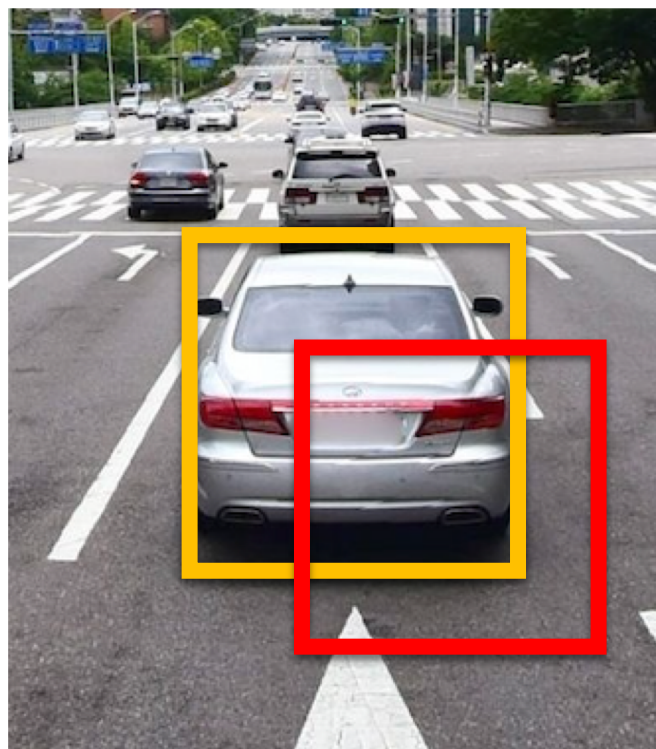
Feature Pyramid Network 구조 [2]

[1] S. Ren et al, Faster R-CNN: Towards real-time object detection with region proposal network, NIPS, 2015

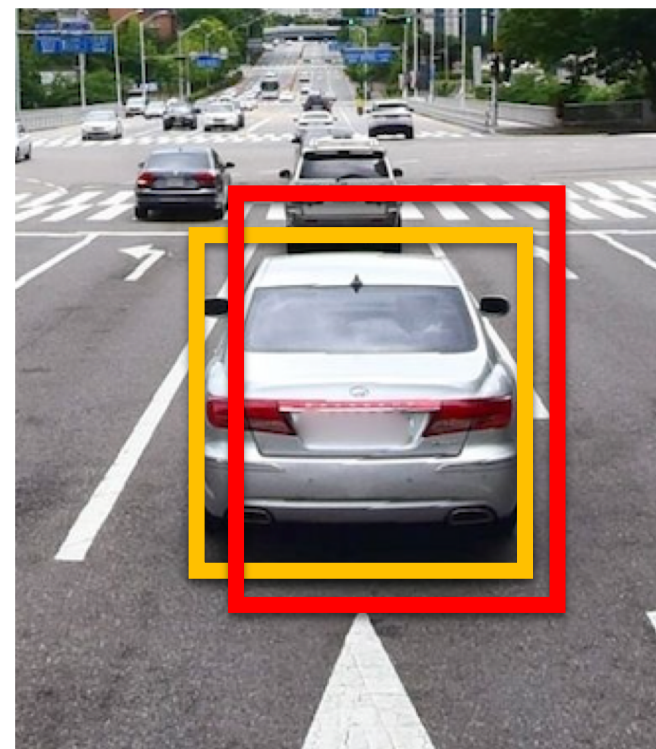
[2] T. Lin et al, Feature Pyramid Networks for Object detection, CVPR, 2017

5.2. Object detection - 평가

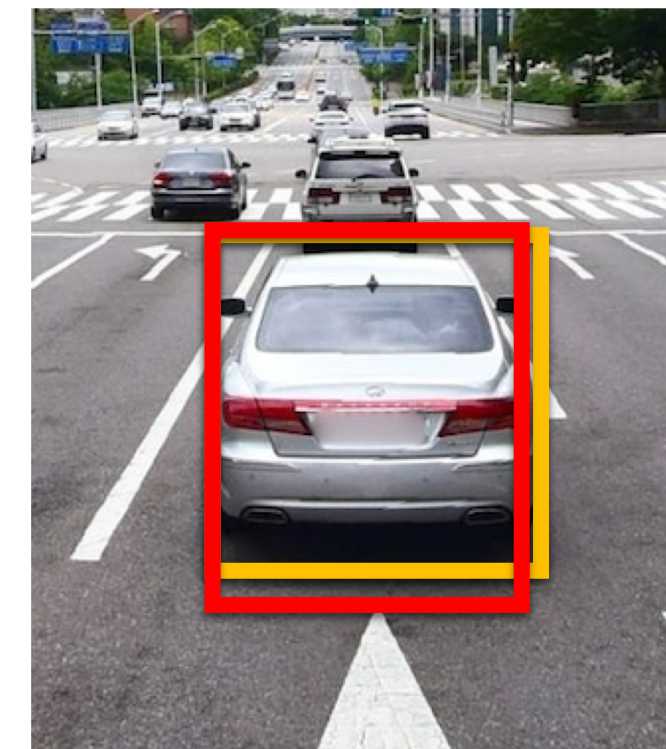
- Object detection 정확도 기준
 - IoU (Intersection over Union) : Bounding box의 예측 정확도
 - AP (Average Precision) : $AP@[0.5:0.95]$
 - mAP (Mean Average Precision) : 각 카테고리 별 AP의 평균
 - AP50 : 각 카테고리 별 $AP@[0.5]$ 의 평균



$IoU = 0.25$



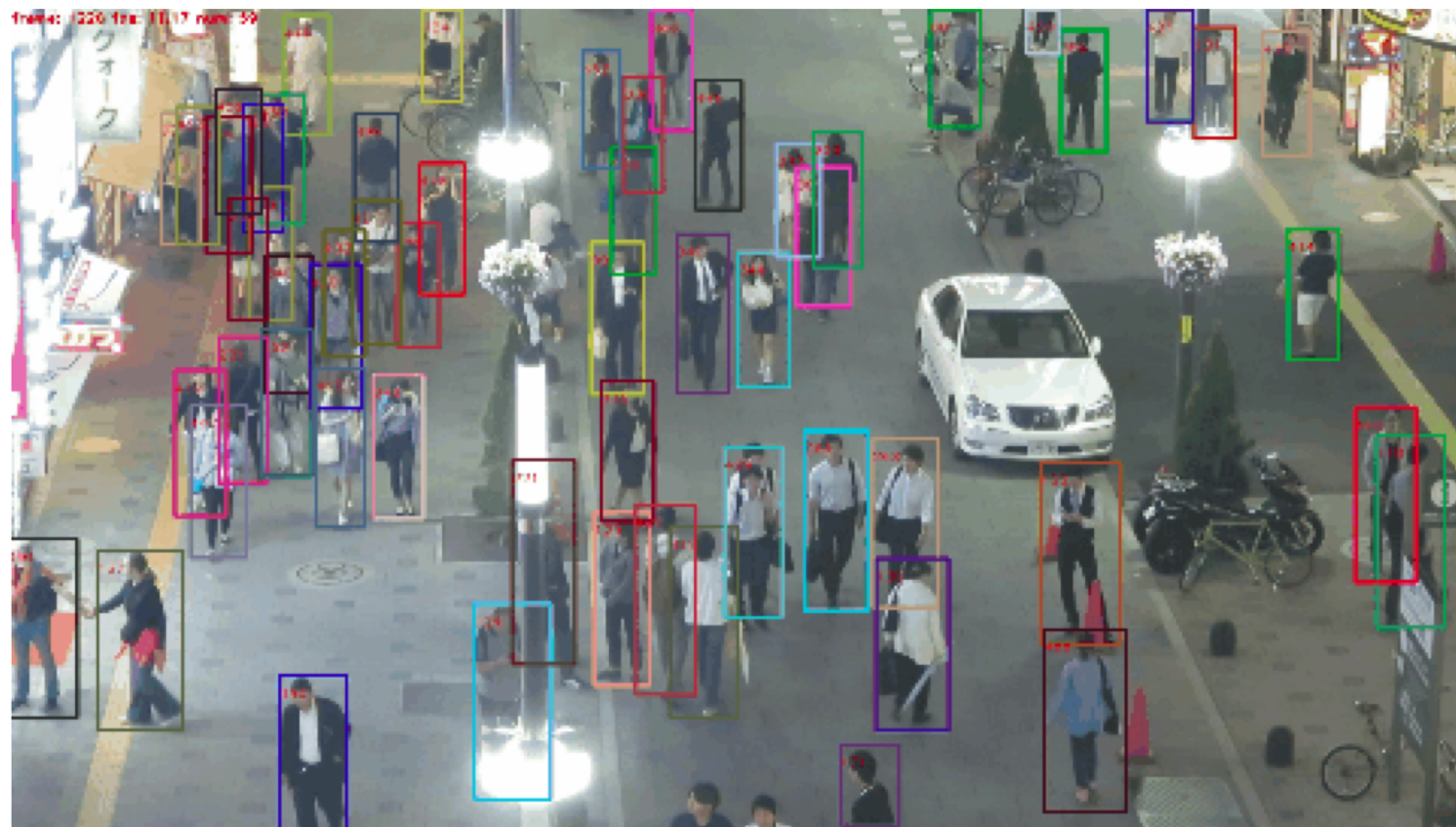
$IoU = 0.5$



$IoU = 0.75$

5.3. Multi-Object Tracking

- Multi-Object Tracking
 - 주어진 연속된 이미지에서 객체만의 고유 ID를 부여
 - 다른 이미지에서 동일한 객체를 구분
 - Motion 정보 (IoU 등 이용) + Appearance 정보 (ReID embedding 이용)



Multi-Object Tracking 예시 | (MOT benchmark) [3]

5.3. Multi Object Tracking – 모델 요구사항

- 서비스/데이터 특성에 따른 모델 요구사항

서비스 특성

- 배치 서비스가 가능한 형태
- 유연한 서빙 환경 (CPU, GPU, 모바일..)

데이터 특성

- 이미지 간 겹치는 부분이 적음
- 다른 객체 간 외양 차이가 큼

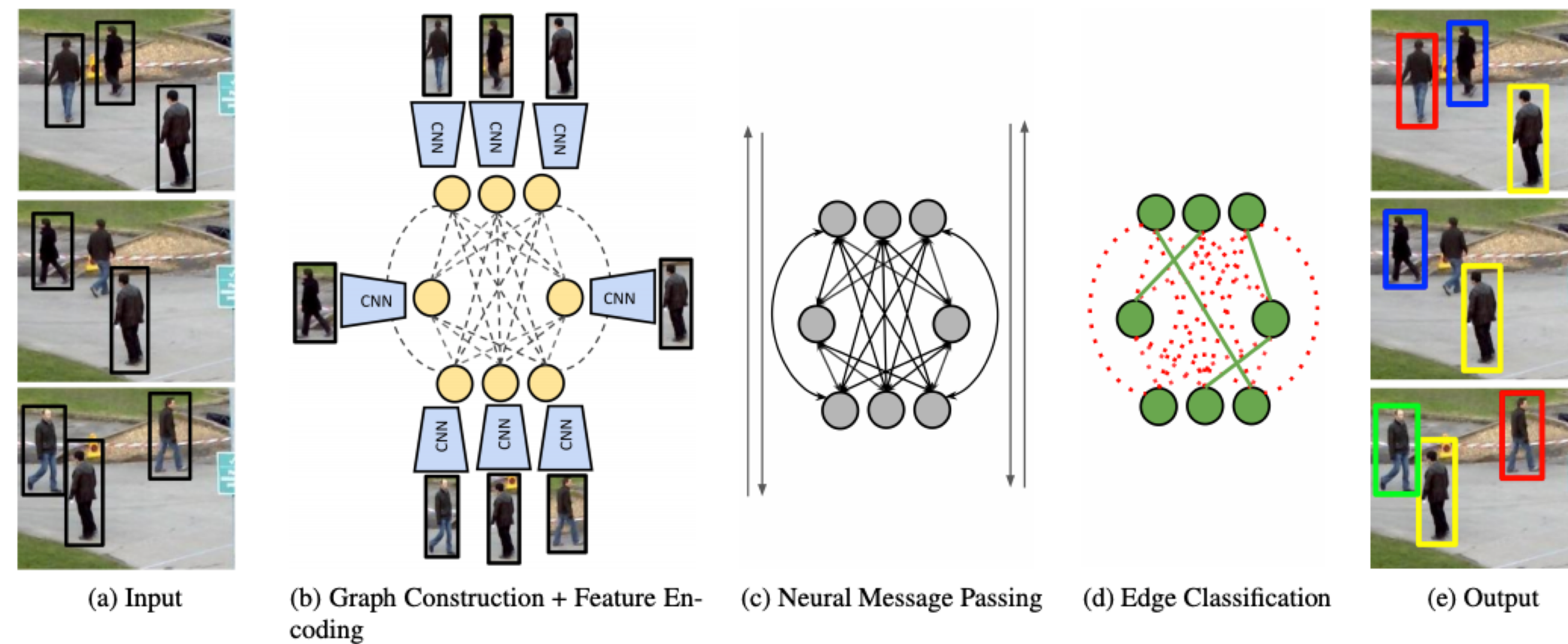


모델 요구사항

- 계산량 절감 < 정확도
- Motion 정보에 IoU 적용 불가
- (Motion + appearance) joint 매칭

5.3. Multi Object Tracking – Baseline 모델

- Message passing network
 - Graph network의 edge = 동일 객체 매칭



Message passing network 적용 예시 [4]

5.3. Multi-Object tracking – 모델 평가

- Multi-Object Tracking 정확도 측정 [5]
 - MOTA : Detection error (FN+FP) & Fragmentation 반영

$$\text{MOTA} = 1 - \frac{FN + FP + \Phi}{T}$$

- IDF1 : Identification precision (IDP)와 Identification recall (IDR)의 F1

$$\text{IDF}_1 = \frac{2\text{IDTP}}{2\text{IDTP} + \text{IDFP} + \text{IDFN}}$$

6. 모델 개발

6.1. 데이터 부족

- 데이터 전처리 및 augmentation 효과
 - 데이터 전처리 및 augmentation 효과

Input size	Data augmentation	Relative AP
(800, 1333)	flip, multi-scale ([640,800]),	1.000
(800, 1333)	flip, multi-scale ([640,800]), rotation([0, 20])	1.026
(800, 1333)	flip, multi-scale ([640,800]), color-jitter ([0.6, 1.4])	1.029
(1024, 1440)	flip, multi-scale ([640,1440]), color-jitter ([0.6, 1.4])	1.231
(1024, 1440)	flip, multi-scale ([640,1440]), color-jitter ([0.6, 1.4]), crop ([0.7, 1.0])	1.237

표. Data augmentation에 따른 시설물 검출 모델 AP 개선

6.1. 데이터 부족

- 데이터 추가에 따른 AP 개선 효과
 - 데이터 추가에 따른 AP 개선 효과
 - 표지판의 경우, # example = 약 1K 부터 AP 개선 미미

# images	Relative AP
5.4K	1
12K	1.041
23K	1.067

표. 데이터 증가에 따른 시설물 검출 모델 AP 개선

6.2. Class imbalance

- Class imbalance
 - 데이터 수가 상대적으로 부족한 category의 경우, 낮은 검출 정확도
 - 학습 instance 수와 classifier weight norm의 관계
 - Balanced group softmax module 제안 [6]

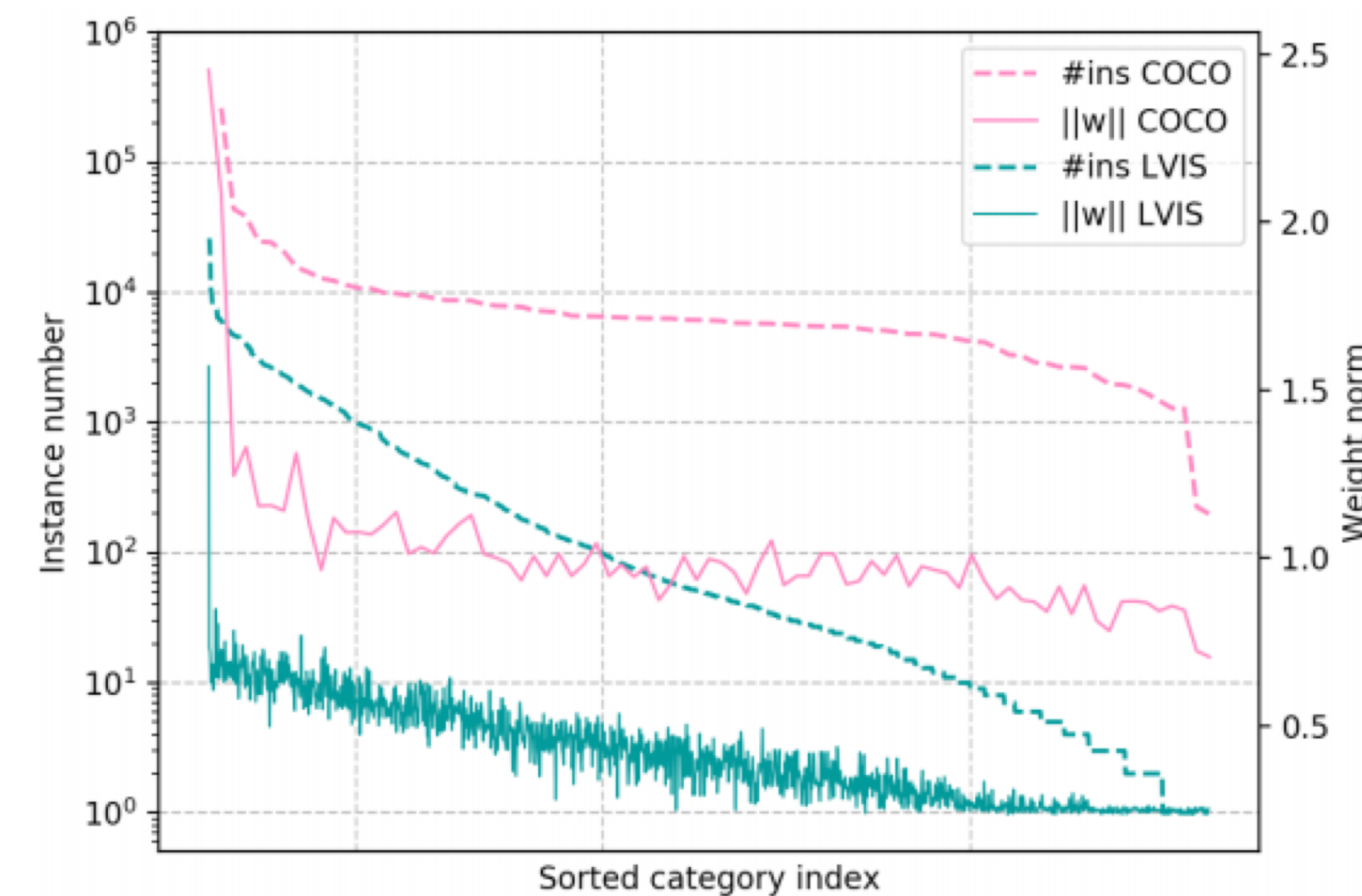


표. 학습 instance 수와 classifier weight norm의 관계

6.2. Class imbalance

- Balanced soft-max module [6] 참고 새로운 method 고안
 - 전체 AP 2% 증가 → 다른 카테고리의 성능 변화 1% 미만
 - 카메라의 Relative AP 약 16% 개선

모델		노면정보	일반표지판	규제표지판	지시표지판	주의표지판	카메라
	# examples	7920	4172	1734	1105	889	153
Baseline	Relative AP	1.482	1.365	1.836	1.700	1.583	1.000
Proposed	Relative AP	1.472	1.351	1.845	1.709	1.591	1.166

표. Class imbalance 개선 모델을 이용한 AP 개선효과

6.3. Sparse image for tracking

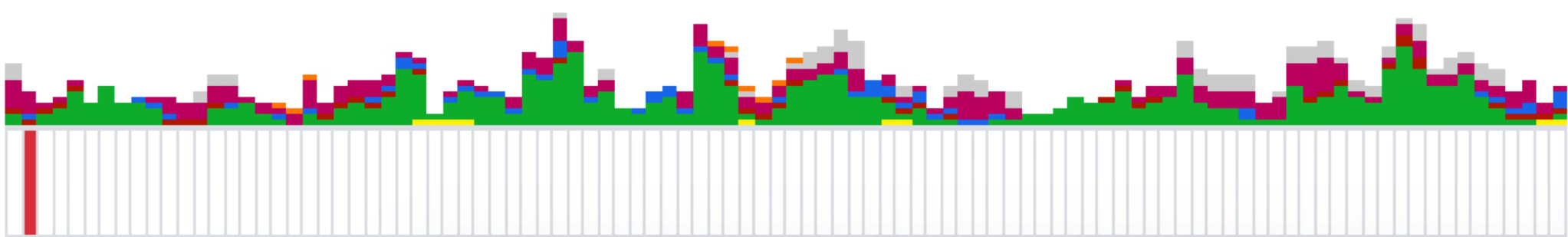
- Bounding box 정보 이용에 대한 ablation study
 - Bbox2node: node에 bounding box 정보 추가
 - Bbox_diff: edge attribute에 bounding box 변화량 추가
 - Bounding box 정보 이용 시 4% Relative IDF1 개선

Bbox2node	Bbox_diff	Relative IDF1
X	X	1
O	X	1.0353
X	O	1.0439
O	O	1.0443

표. Bounding box 정보 추가에 따른 IDF1 개선 효과

141 (덕릉로) ▼

19159001286



v1.0.1 ▼

Summary

Total Sign Count	165
단속카메라	2개

차선정보

차선번호	1차선	2차선	3차선	4차선
방향 방면				

시설물 정보

category	시설물/안전운행	조건/상세 정보	Tracking (DB갱신시점)
지시	자전거 및 보행자 통행 구분		DB업데이트 요청
지시	견인지역		DB업데이트 요청

7. Future work

다음에는 무엇을 시도해볼까?

전면의 거리뷰를 활용했으니...

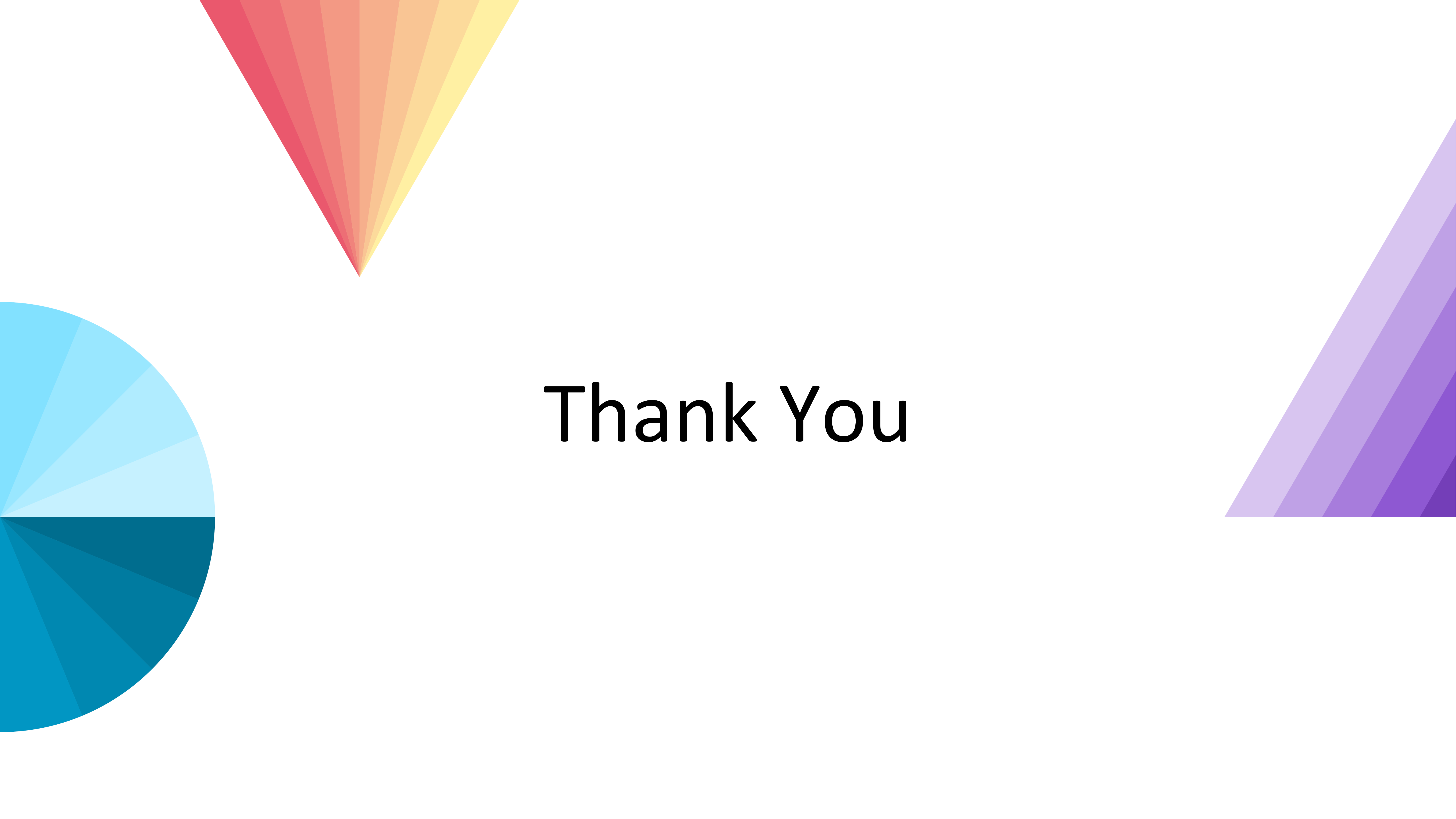




Q & A

Reference

- [1] S. Ren et al, Faster R-CNN: Towards real-time object detection with region proposal network, NIPS, 2015
- [2] T. Lin et al, Feature Pyramid Networks for Object detection, CVPR, 2017
- [3] Zhongdao, Towards-Realtime-MOT Github, <https://github.com/Zhongdao/Towards-Realtime-MOT>
- [4] G. Braso et al, Learning a neural solver for multiple object tracking, *CVPR*, 2020
- [5] E. Ristani et al, Performance measures and a data set for MTMC tracking, In *ECCVW*, 2016



Thank You