



KOREA
UNIVERSITY

INISW Academy
지능정보SW아카데미

실종 아동 트래킹 기반 지능형 CCTV 솔루션

7조 백대환, 손태정, 송보미, 윤연경



목차



1

서비스 구상 배경

문제 인식
해결 방안

2

모델 소개

사용 모델 소개
모델의 필요성

3

파인튜닝 과정

4

모델 파이프라인

모델 파이프라인
트래킹 작동 방식

5

모델 시연

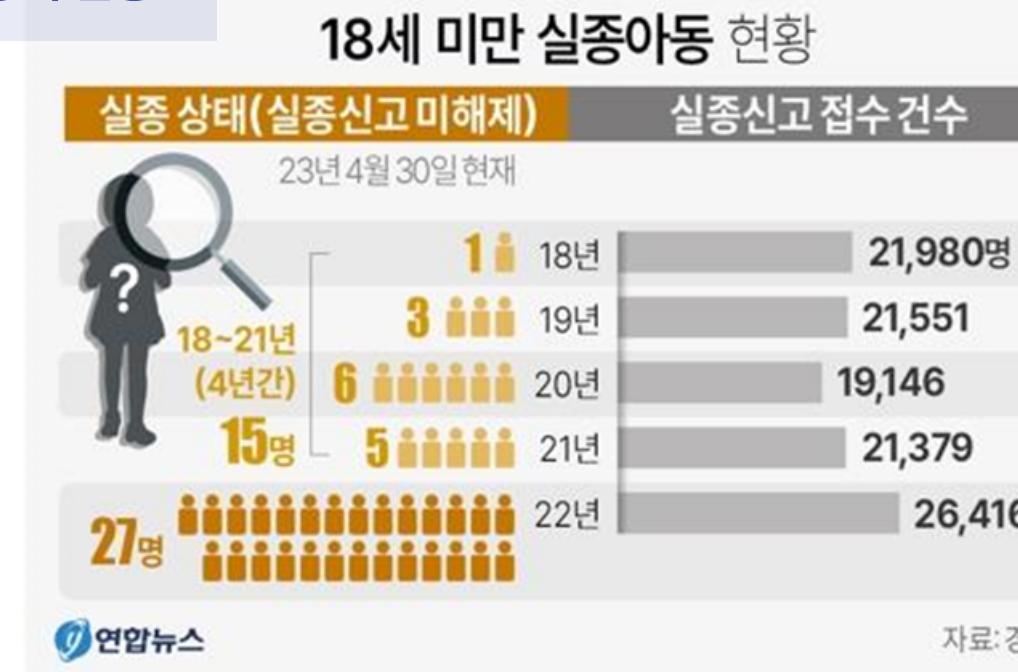
6

기대효과 및 활용방안

서비스 구상 배경

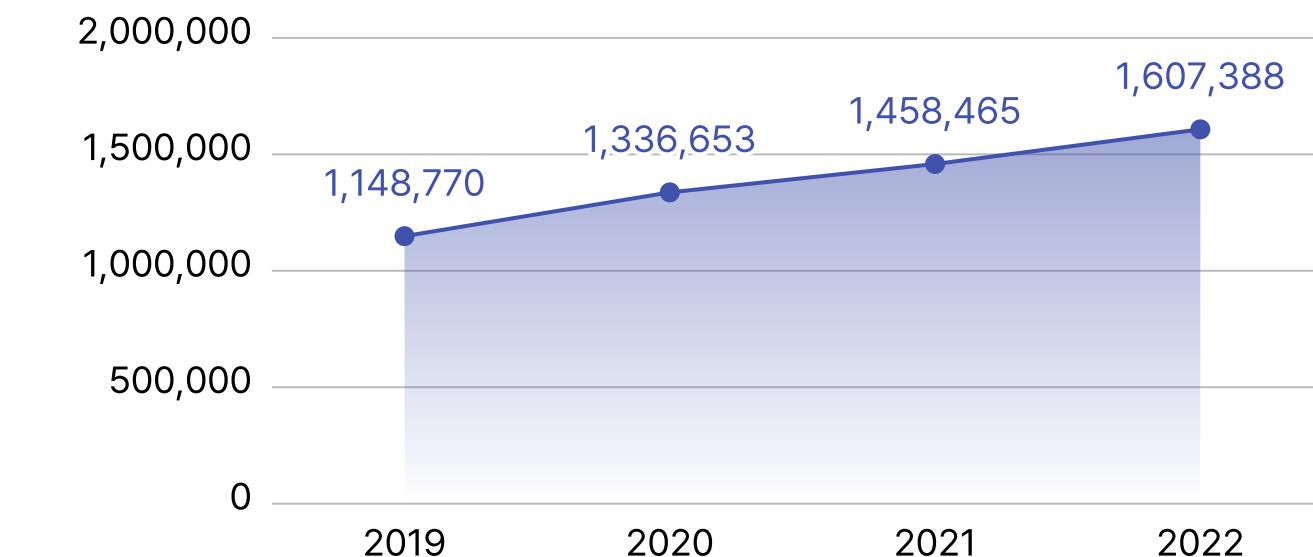
문제 인식

최근 5년 간 실종아동 상태 현황



최근 4년 간 CCTV 설치 대수

자료 : 개인정보보호 종합지원시스템



현재 CCTV 통합 관제 시스템



비효율적인 추적 방식

시도	관제센터 수 (개소)	관제인력(명)				통합인기 CCTV (대수)	1인당 CCTV (대수)
		계	지자체	경찰	위탁		
서울	25	491	184	100	195	12	75,846
부산	10	331	248	48	33	2	21,675
대구	8	239	220	10	-	-	12,421
인천	11	220	100	33	71	16	16,975
광주	1	93	89	33	-	-	8,013
대전	1	66	61	5	-	-	5,421
울산	5	101	1	15	84	1	9,668

한정적인 운용 인력

문제 인식

최근 5년 간 실종아동 상태 현황

18세 미만 실종아동 현황



최근 4년 간 CCTV 설치 대수

자료 : 개인정보보호 종합지원시스템



인물 추적 및 동선 파악의 효율성을 위한 자동 추적 시스템 필요!

CCTV 통합 관제 시스템

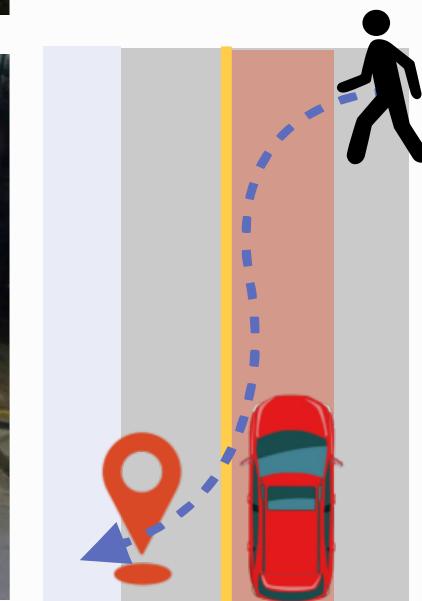
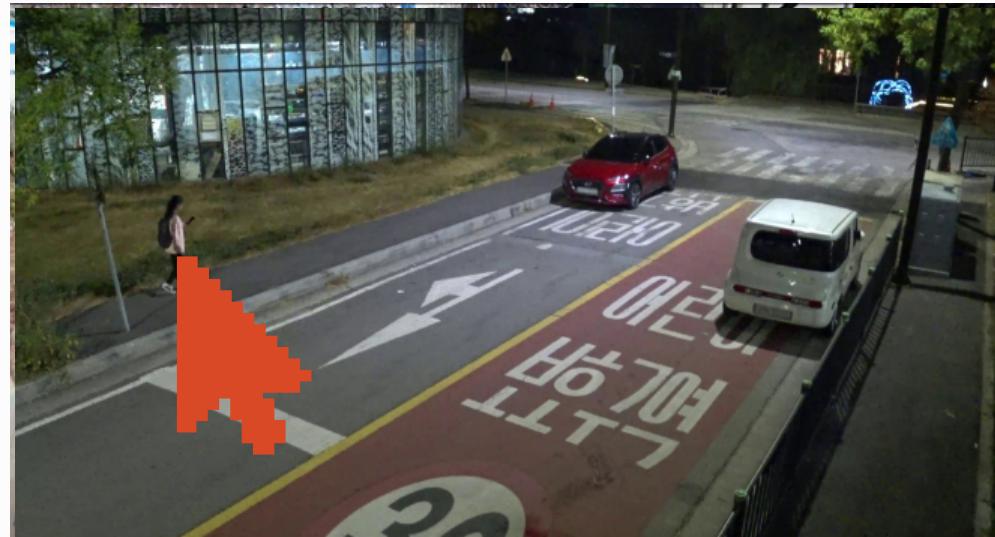


시도	단위 (개소)	계	지자체	경찰			CCTV (대수)	CCTV (대수)
				위탁	공익	경찰		
서울	25	491	184	100	195	12	75,846	154
부산	10	331	248	48	33	2	21,675	65
대구	8	239	220	10	-	-	12,421	52
인천	11	220	100	33	71	16	16,975	77
광주	1	93	89	33	-	-	8,013	86
대전	1	66	61	5	-	-	5,421	82
울산	5	101	1	15	84	1	9,668	95

한정적인 운용 인력

해결 방안

사용자가 직접
추적할 대상 설정



지능형 cctv의 장점

01 / CCTV의 효율적 사용 필요성

전국적으로 CCTV의 수는 늘어나는 데에 비해
연간 실종 아동 수는 증가하는 추세. CCTV의 확대를
효율적으로 이용해 운용인력, 비용 등을 절감 가능

02 / 실종 아동 발견시간 단축의 중요성

실종 신고 후 12시간이 지나게 되면 못 찾을 확률이
58%, 24시간이 지나면 68%, 그리고 일주일이 지나면
90%까지 올라간다는 경찰청 통계 존재

03 / 객체 자동 추적의 효과성

객체를 자동으로 추적하면 경로 파악에 용이하여 실종
아동의 위치 추적 뿐 아니라 범죄 추적 등에도 이용될 수
있음. 노숙자나 주취자의 안전 사고 등도 미연에 방지

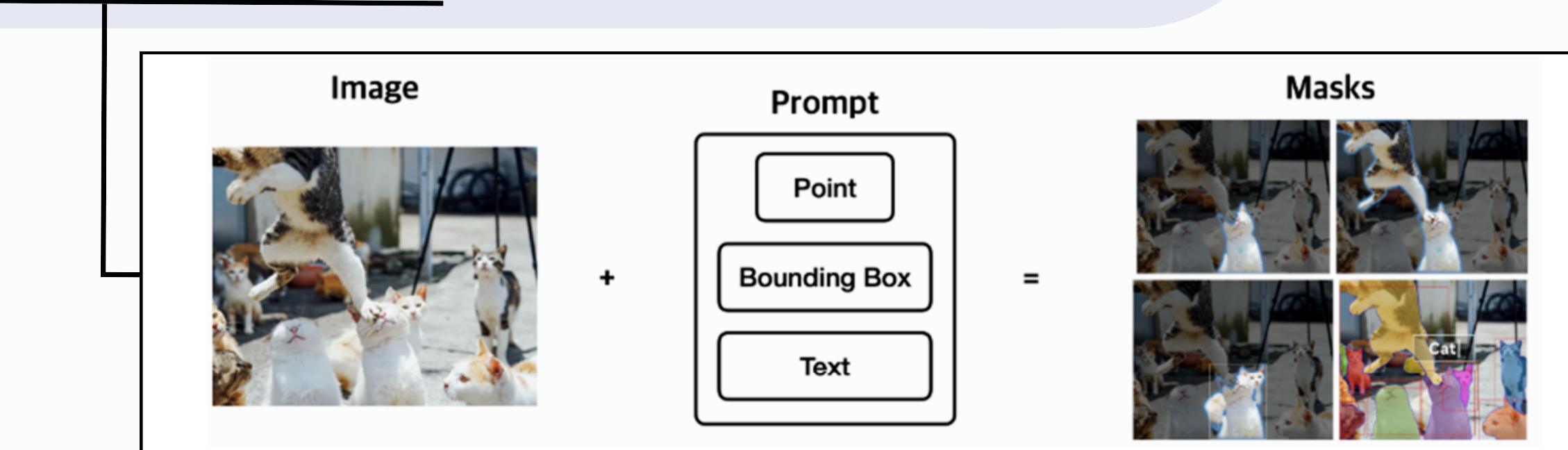
모델 소개

사용 모델 소개

Segment Anything Model (SAM)

올해 초 메타에서 발표한 image segmentation model로
대용량 데이터셋(SA-1B) 학습으로 모든 분야에서 광범위하게
사용할 수 있는 foundation model

추가적인 훈련 없이 pre-trained 모델을 새로운 task에
적용할 수 있는 Zero-shot generalization 능력과
점, 박스, 텍스트 프롬프트 입력을 통해 마스크 생성 가능



사용 모델 소개

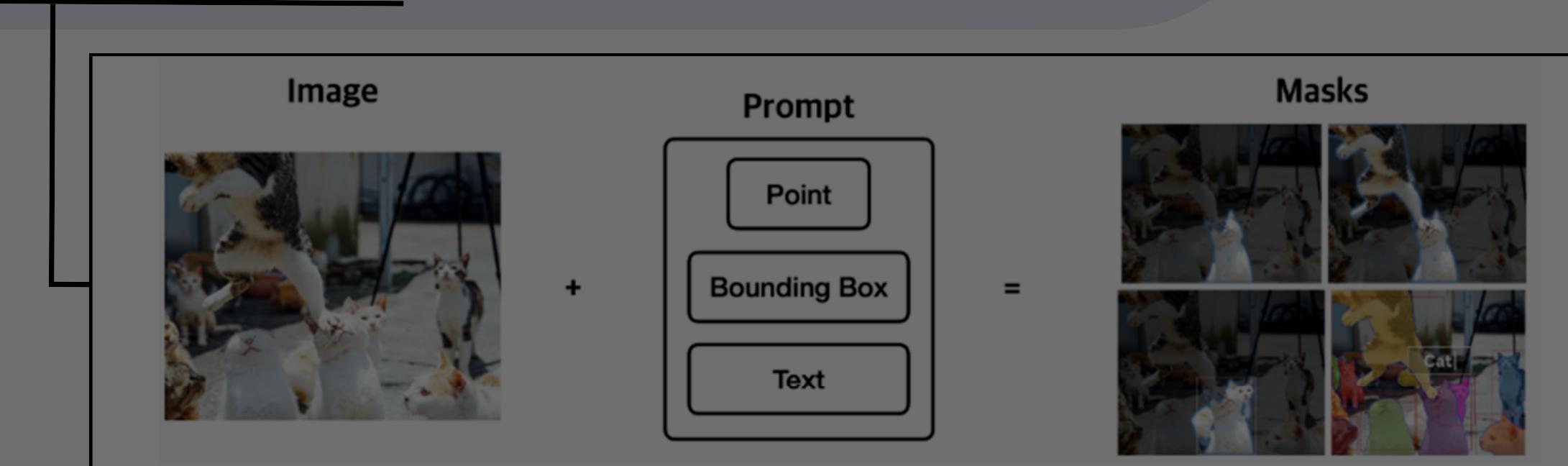
Segment Anything Model

올해 초 메타에서 발표한 image segmentation model로
대용량 데이터셋(SA-1B) 학습으로 모든 분야에서 광범위하게
사용할 수 있는 foundation model



많은 양의 CCTV 데이터에서 미아만을 분리해내기 적절할 것이라고 판단

점, 박스, 텍스트 프롬프트 입력을 통해 마스크 생성 가능



사용 모델 소개

SAM

segmentation된
mask 제공



Xmem

이전 프레임을 기반으로
다음 프레임의 객체 mask를
예측해 tracking 진행

tracking에서 발생하는 겹침 현상 해소

사용자의 입력을 통한 interactive한 방식

zero-shot segmentation으로
많은 양의 데이터 학습 부담 감소

GPU 효율적으로 사용해
긴 영상 처리에 효과적

사용 모델 소개



vs



경량화 모델보다
segmentation 정확도 ↑



└ SAM의 경량화 모델 ┘

모델의 필요성

Previous

- 기존의 트래킹은 yolo detection을 이용하여 인식되는 모든 객체를 추적하는 방식
- 동일 객체를 제대로 tracking 하지 못하고, tracking하고 싶은 객체를 눈으로 쫓아가야 했음

Ours

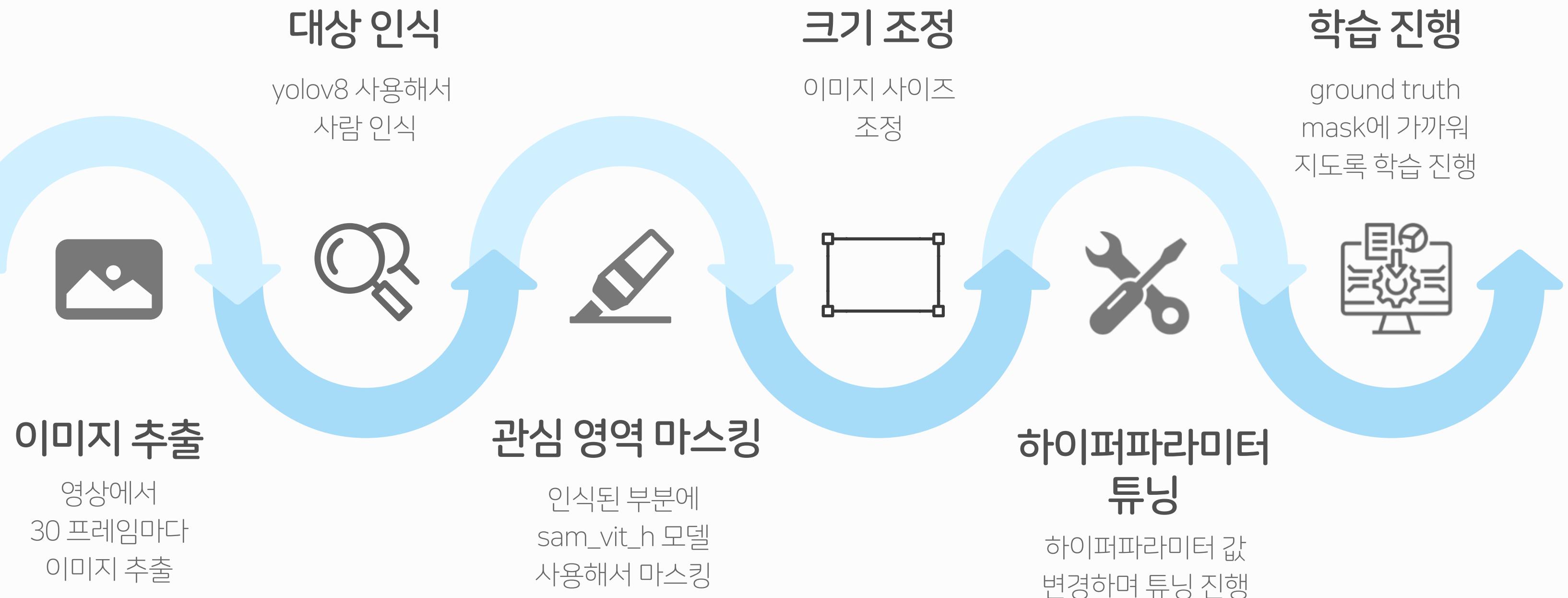
- SAM 모델의 경량화를 통해 빠른 segmentaiton 가능
- 사용자와의 interaction을 통해 관심 있는 객체만 트래킹함으로써 효율성 극대화



cctv와 같은 multi camera 환경에서 빠르게 관심 대상을 추적할 수 있음

파인튜닝 과정

파인튜닝 과정

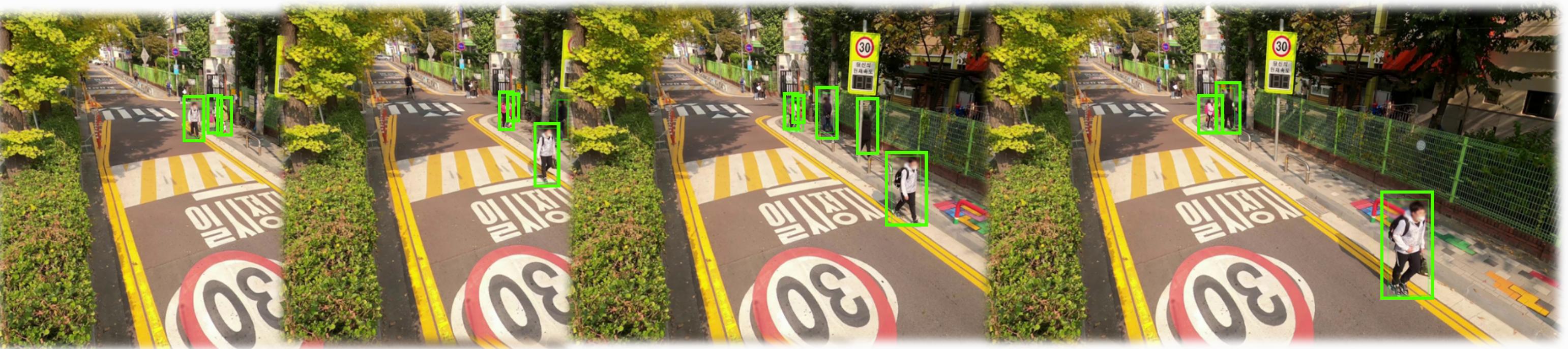


파인튜닝 과정

이미지 추출



대상 인식

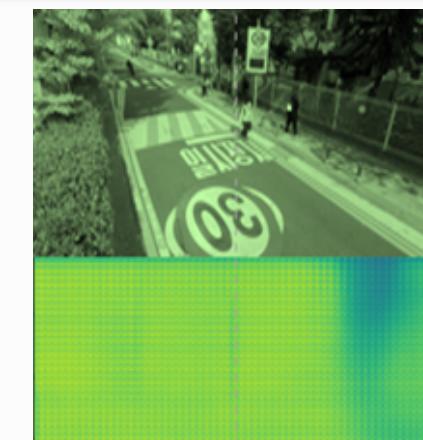


파인튜닝 과정

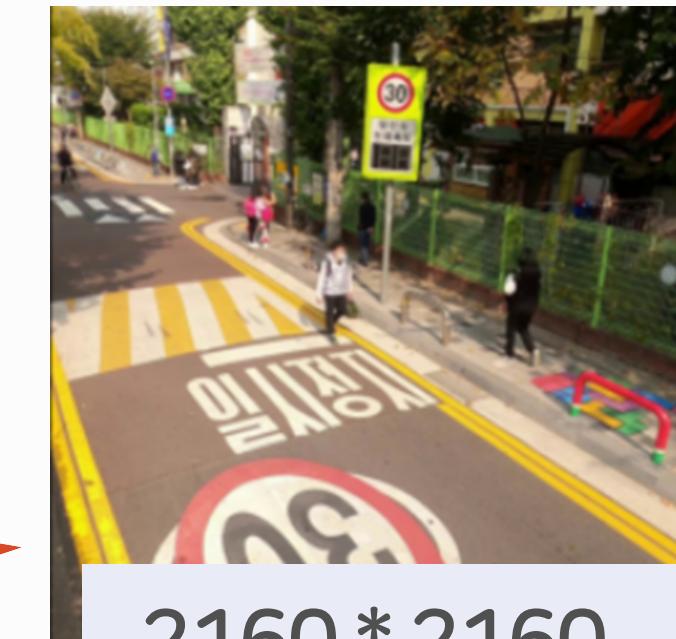
관심 영역 마스킹



크기 조정



.....>
오류 발생



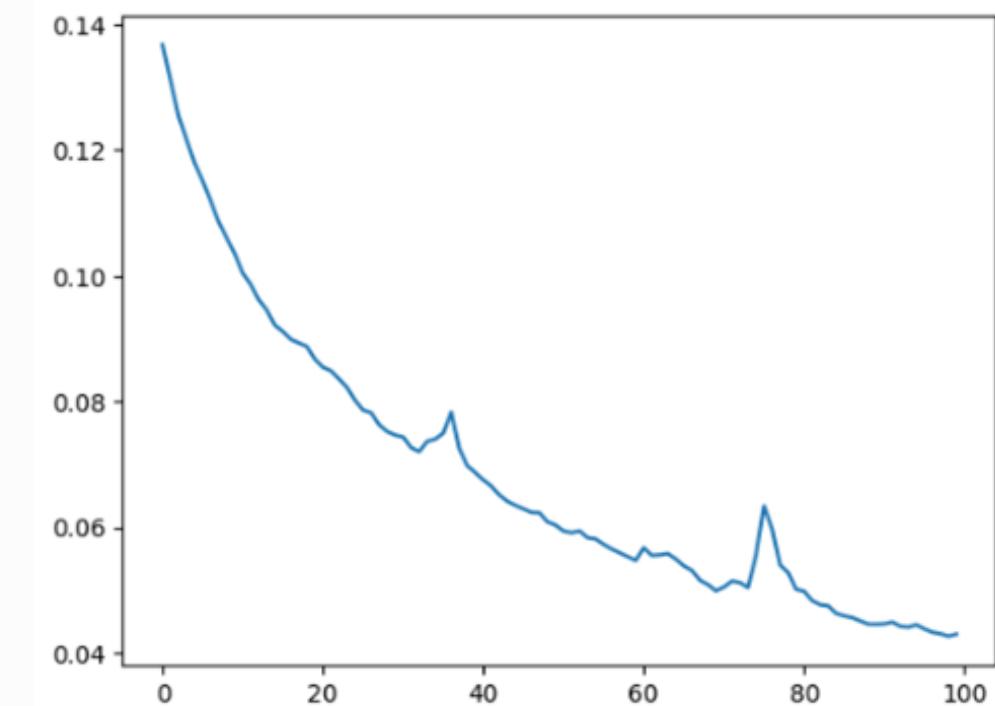
.....>
모델에 적합한
사이즈로 변경



파인튜닝 과정

하이퍼파라미터 튜닝

파인튜닝 위한 이미지 구조인 $N \times N$ 형태로 crop해서 인물이 찍히는 사진들로 파인튜닝 진행. 데이터 전처리 후, 약 2000개의 데이터로 분석 진행. Loss는 DiceCELoss 사용. Optimizer는 Adam 사용하여 하이퍼파라미터 튜닝.



학습 진행

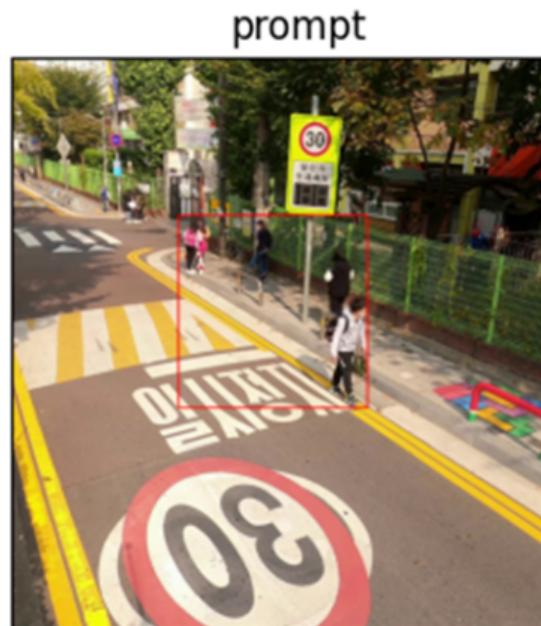
이후 새로운 사진으로 predict를 진행.
높은 IoU값을 보임.

파인튜닝 과정

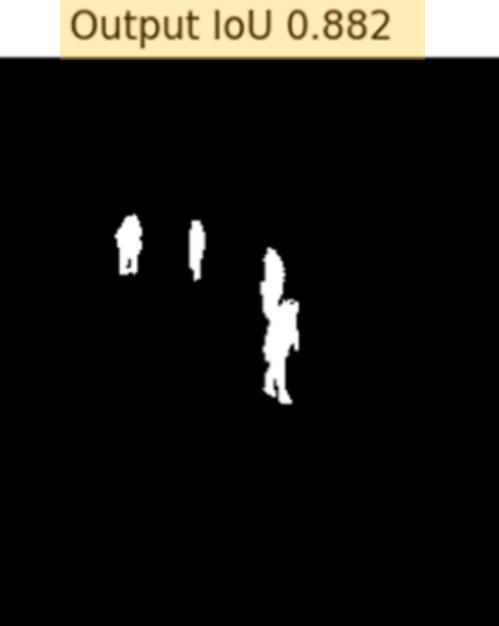
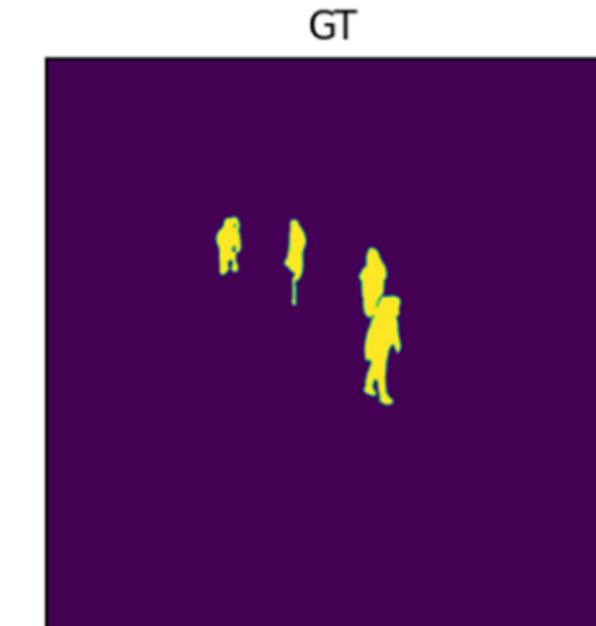
원본



prompt 입력



ground truth와 학습 결과 비교



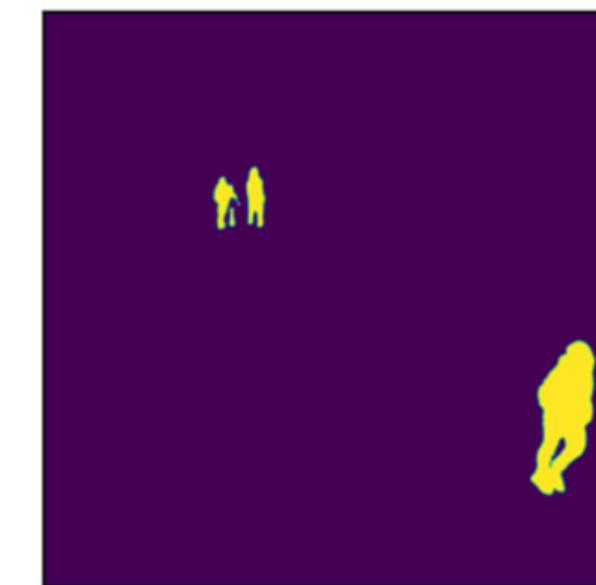
Input_Image



prompt



GT

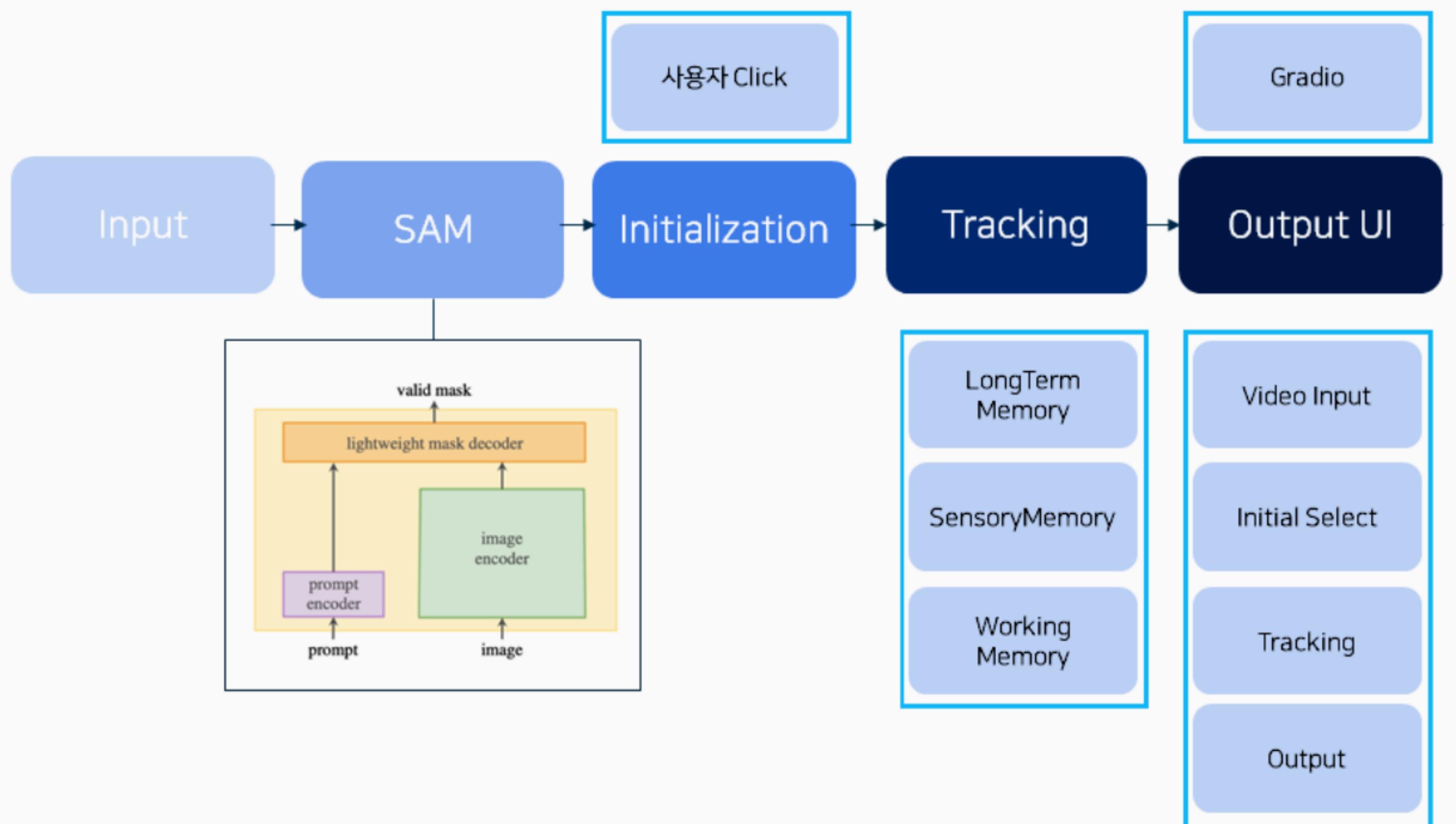


Output IoU 0.895

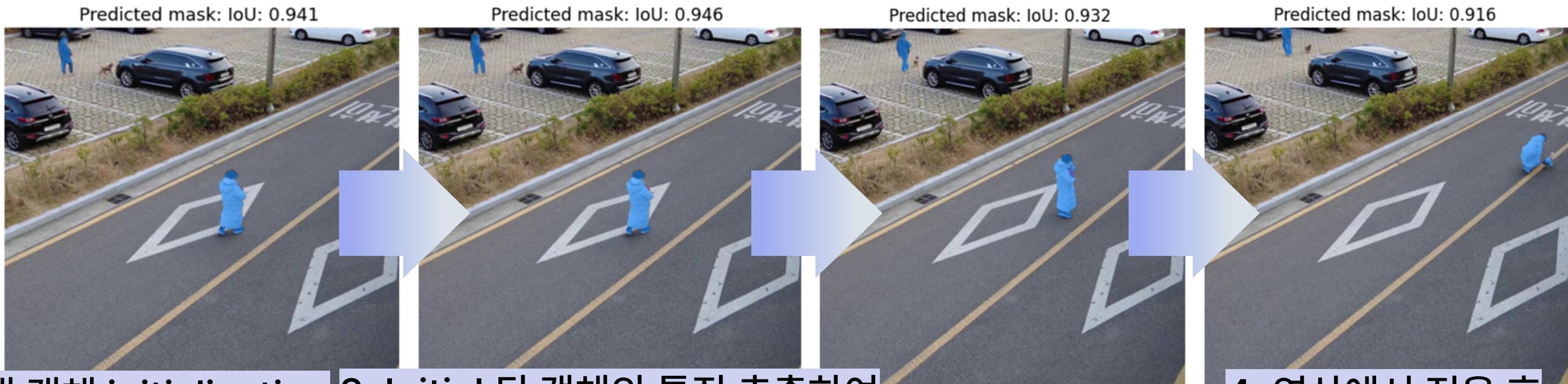


모델 파이프라인

모델 파이프라인



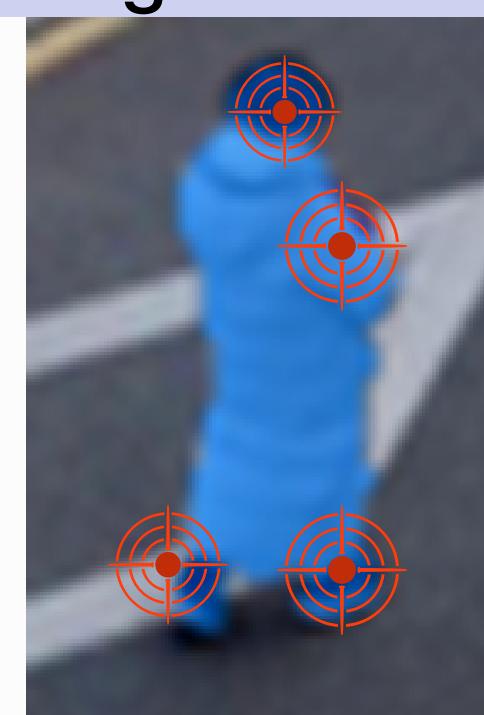
트래킹 작동 방식



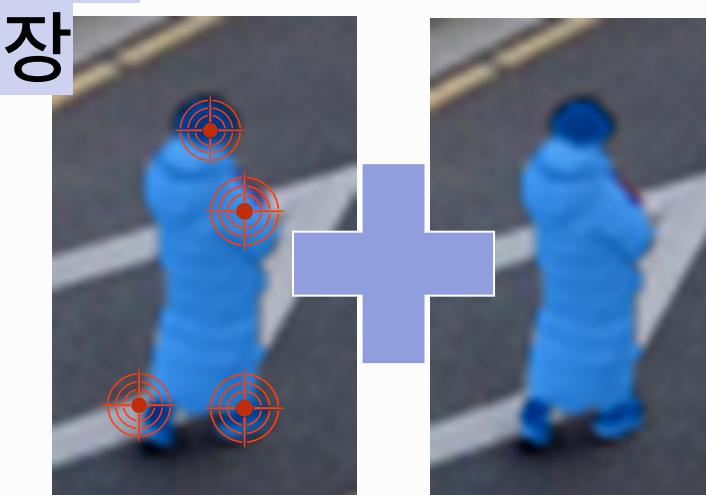
1. 영상 내 객체 initialization



2. Initial 된 객체의 특징 추출하여
Long term memory에 저장



4. 영상에서 적용 후
Output 확인



3. sensory memory의 객체를 학습,
Long term memory의 값들과
결합하여 정확도 Upgrade

모델 시연

기대효과 및 활용방안

기대효과 및 활용방안



다중 CCTV 영상을 통한 사람 탐지, 재인식, 추적 서비스

1

인물 추적
솔루션 제공

2

고객 행동 분석
솔루션 제공

3

동물 활동 반경 분석
솔루션 제공



다양한 장소에서 범죄 용의자, 미아 등
특정 객체에 대한 탐지 및 추적을
할 수 있음



인물 추적으로 고객 유형 간 행동 패턴
비교, 분석 등 맞춤형 고객 행동 분석
가능, 골목 상권 분석 등에도 용이함

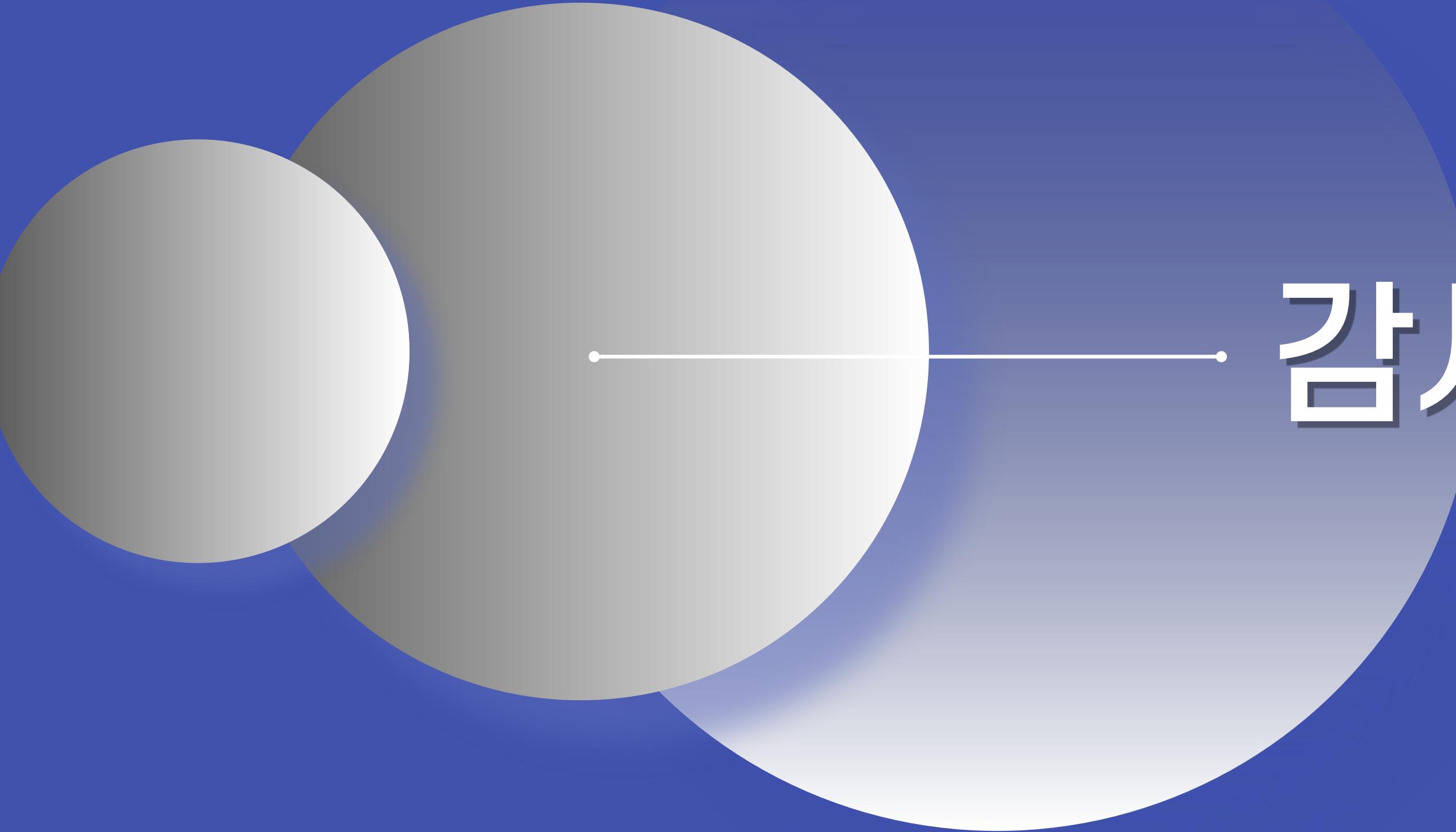


야생동물, 유기견, 천연기념물 등 특정
동물 탐지 및 추적을 통해 활동 범위와
활동 분포 분석, 동물 보호에 용이함



INISW Academy

지능정보SW아카데미



감사합니다!