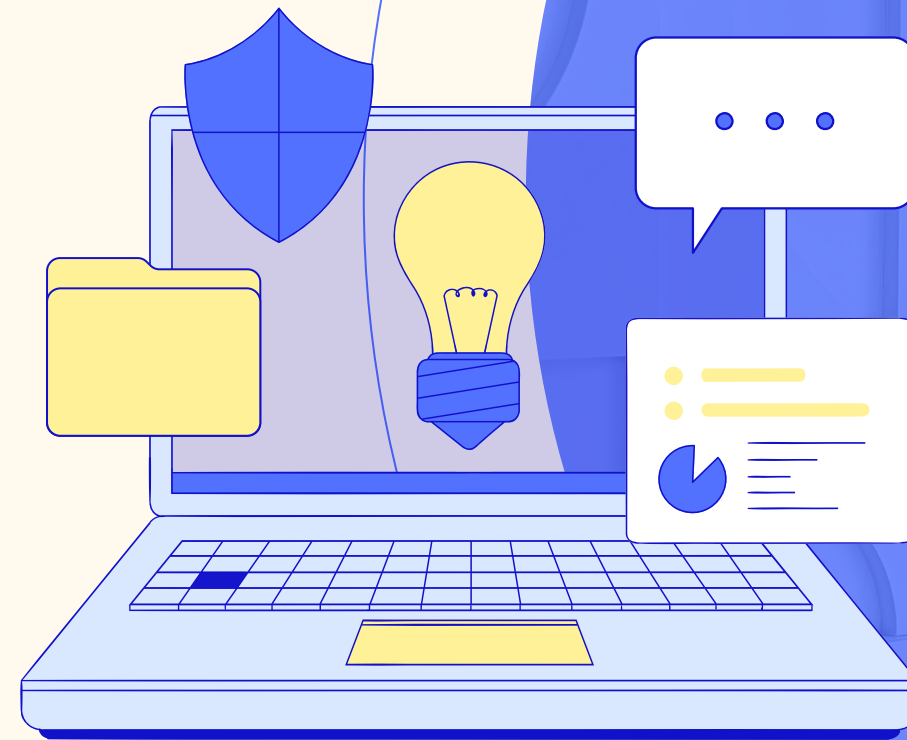


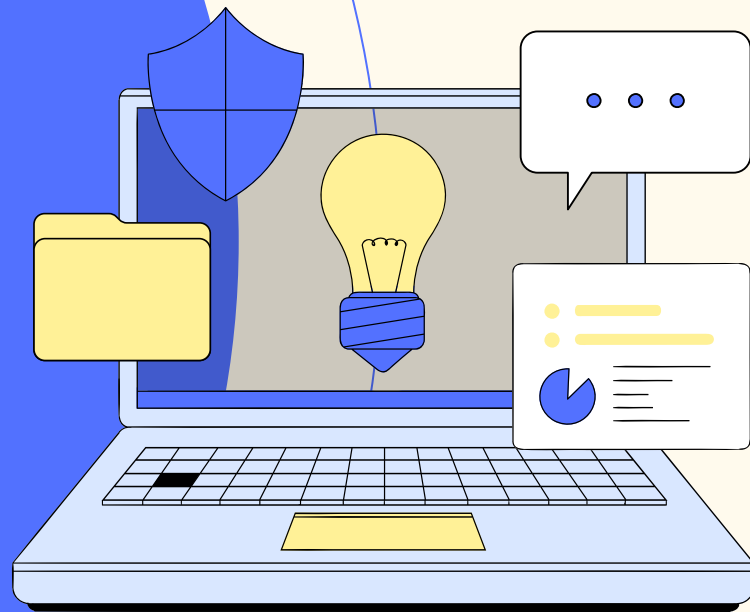
시각장애인을 위한 스마트폰 사진기 음성안내 서비스

CV2조

12기 김지원 박석우
13기 박소연 배성윤



목차



01

프로젝트 주제 소개

02

진행상황 TimeLine

03

향후 계획

CONTENTS

프로젝트 주제 소개 ●

진행상황
TimeLine

향후 계획

01

프로젝트 주제 소개

“시각장애인을 위한 스마트폰 사진기 음성안내 서비스”

- 서비스 대상 : 시각장애인
- 서비스 목표 : 시각장애인 이용자가 더욱 수월하게 사진을 촬영할 수 있도록 음성으로 안내하는 서비스 제공
- 프로젝트 목표 :
 - A. REAL-TIME OBJECT DETECTION MODEL을 서비스에 맞는 FACE DETECTION MODEL로 전이학습
 - B. 모델 추론의 정확도 향상 및 알고리즘 최적화를 통한 실시간 추론 속도 개선
 - C. STREAMLIT 웹 구현

주제 : 시각장애인을 위한 스마트폰 사진기 음성안내 서비스

서비스 대상

- 시각장애인

서비스 목표

- 사용자가 상황별로 사진을 찍을 때 사진을 더 편리하게 찍을 수 있도록 음성으로 안내하는 서비스를 제공한다.
- opencv와 pretrained object detection model을 이용하여 우리의 task에 맞도록 전이학습시킨다.
- 모델의 학습을 통해 성능 및 정확도를 향상시킨다.
- 알고리즘의 최적화를 통해 latency를 최소화한다.

사용자가 사진을 찍을 때까지

Iteration

출력 : 텍스트

출력 변환 : 음성

카메라의 객체 인식 과정

전면카메라

Face Detection Model (Real-Time Object Detection)

요소1 : 사용자의 얼굴이 프레임에서 차지하는 비중

요소2 : 프레임 중앙 기준 얼굴의 위치

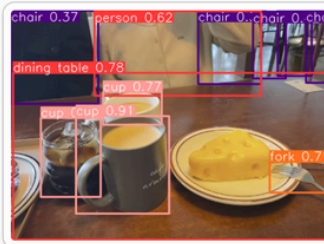
MediaPipe

요소3 : 얼굴의 각도 (바라보고 있는 방향)

후면카메라

인식요소

- 상황 : 피사체가 있는 상황에서



인식된 객체의 label 분류

confidence >= 0.6

객체 1 : 위치
객체 2 : 위치
객체 3 : 위치
...

사용자가
카메라를 움직임

IF "객체 1 : 위치"
(의 상황이 왔을 때)
(+ 객체가 프레임 속에서 차지하는 비율)

if: 얼굴 비중 N% 이상

else:
안내하지 않는다

if: 얼굴 각도 N도 이상

else:
안내하지 않는다

음성안내

- 3요소에 대해 안내 멘트를 발화한다.
- 조건에 따라 1,3요소는 안내되지 않을 수 있다.
- 아래에 나열된 순서대로 발화된다.

"현재 얼굴이 화면의 {비중}% 이상 차지합니다."

"현재 얼굴이 화면의 {구획명} 방향에 있습니다."

"현재 얼굴이 {좌/우}, {상/하} 방향으로
{각도}도 돌아가 있습니다."

최종 구도
확정

사진을
찍는다

CONTENTS

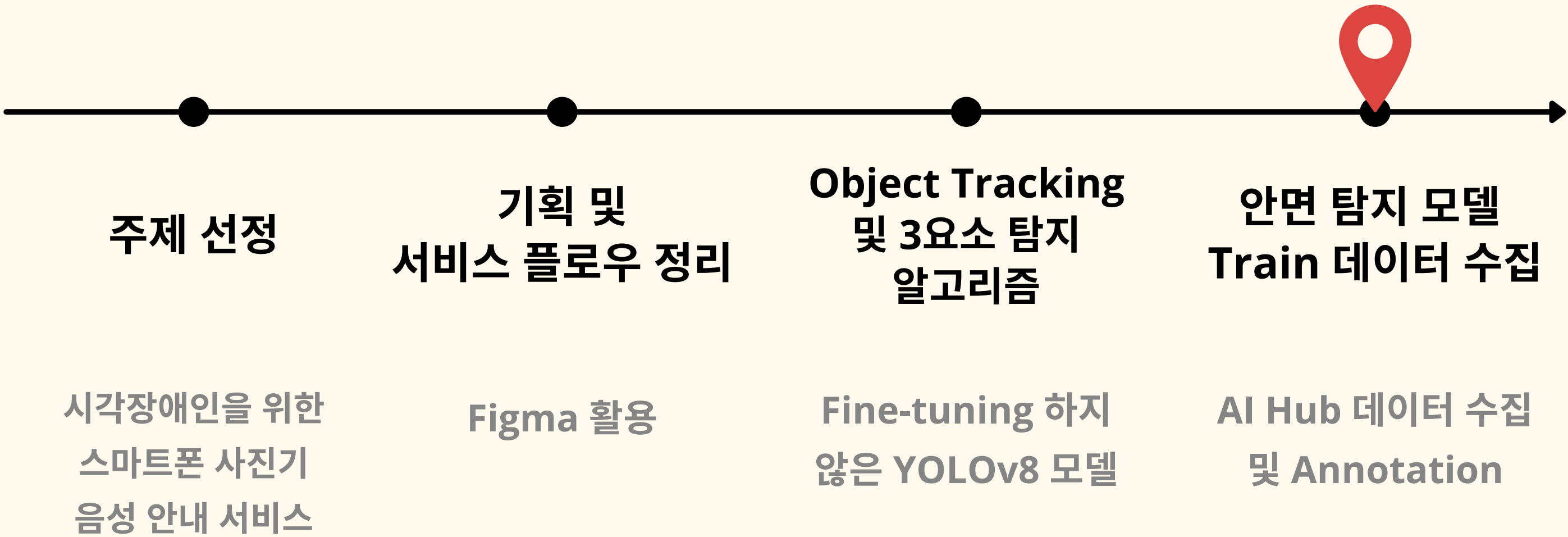
프로젝트 주제 소개

진행상황
TimeLine

향후 계획

02

진행상황 TimeLine



02

진행상황 TimeLine

데이터 및 모델 소개

- 데이터 출처: AI hub

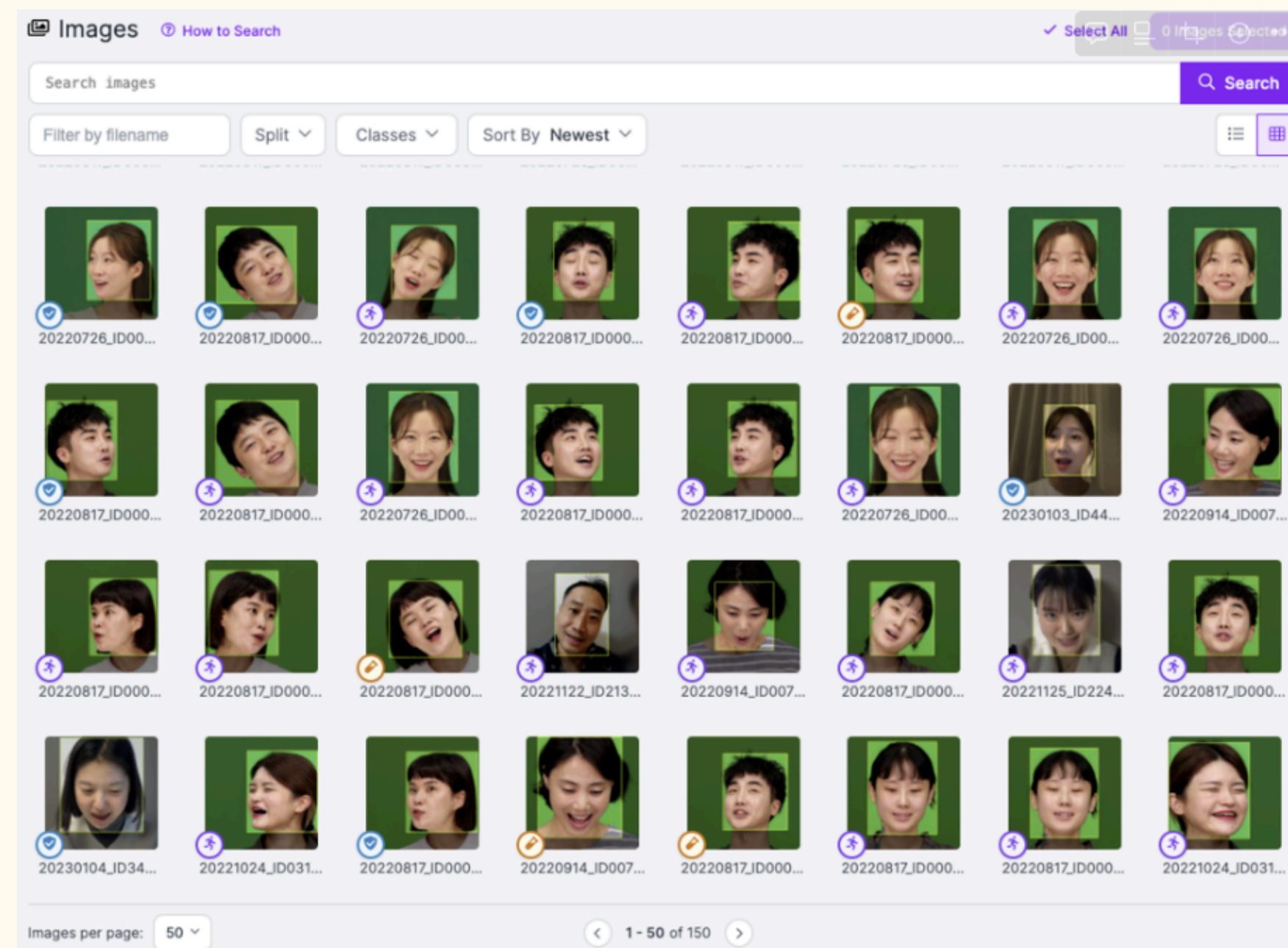
데이터 형식: .png, .txt

다양한 표정과 각도의 얼굴 데이터셋을
annotation하여

fine-tuning에 사용할 데이터 생성

→ 총 1500여개의 데이터 생성 및 활용 예정

- 사용 모델: YOLOv8,
MobileNet + SSD



CONTENTS

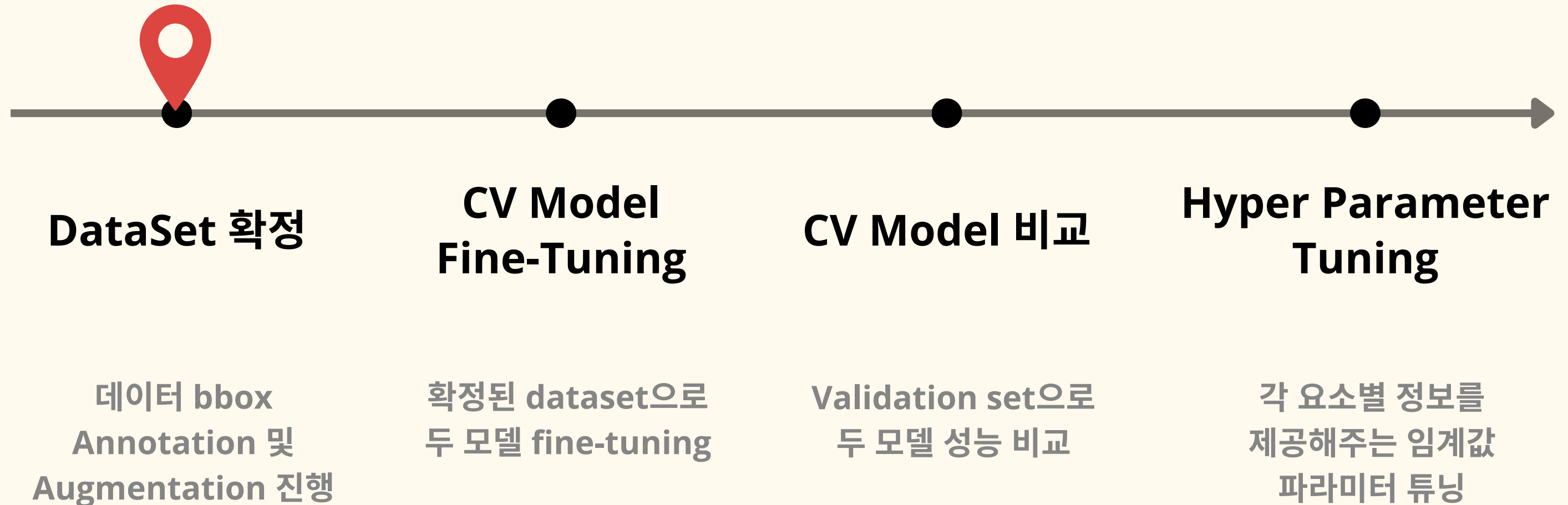
프로젝트 주제 소개

진행상황
TimeLine

향후 계획

03

향후 계획





감사합니다