

# Amazon SageMaker Ground Truth

## 1. Create workforce

**Amazon SageMaker**

Amazon SageMaker Studio

대시보드

검색

Ground Truth

- 레이블링 작업
- 데이터 세트에 레이블 지정
- 레이블링 인력

노트북

- 노트북 인스턴스
- 수명 주기 구성
- Git 리포지토리

처리 중

- 작업 처리 중

훈련

- 알고리즘
- 훈련 작업
- 하이퍼파라미터 튜닝 작업

추론

- 컴파일 작업
- 모델 패키지
- 모델

의견 한국어

Amazon SageMaker > 레이블 지정 작업 인력 > 프라이빗 팀 만들기

**프라이빗 팀 만들기**

**Private team creation**

Create a private team with AWS Cognito  
Create a private work team by sending email invitations to new workers or importing workers from existing Amazon Cognito user groups.

Create a private team with OpenID Connect (OIDC)  
Create a private work team with your own identity provider (IdP). Your IdP must support OIDC user groups.

**팀 세부 정보**

팀 이름  
작업팀에 알아보기 쉬운 이름을 지정합니다. 이 이름은 나중에 변경할 수 없습니다.

최대 63자의 알파벳 문자 및 숫자, 하이픈을 포함할 수 있지만 공백은 포함할 수 없습니다. AWS 리전의 계정에서 고유해야 합니다.

**작업자 추가 정보**

작업자 이메일 주소를 추가하거나, 기존 Amazon Cognito 사용자 그룹에서 작업자를 가져와서 프라이빗 작업팀에 작업자를 추가합니다.

이메일로 새 작업자 초대  
기존 Amazon Cognito 사용자 그룹에서 작업자 가져오기

**이메일 주소**

초대장 미리 보기

© 2008 - 2020, Amazon Web Services, Inc. 또는 계열사. All rights reserved. 개인 정보 보호 정책 이용 약관

- SageMaker의 Labeling workforce에서 Create private team을 선택해 줍니다.
- 각 항목들을 입력하여 private 팀을 생성해 줍니다.

**프라이빗 작업 인력 요약 정보**

작업 인력을 Amazon Cognito 사용자 풀에서 지원합니다. 사용자 풀을 사용하여 작업자는 Amazon Cognito를 통해 레이블 지정 포털에 로그인할 수 있습니다.

Amazon Cognito 사용자 풀  
us-east-1\_8OqZg77p8 [ ]

레이블 지정 포털 로그인 URL  
<https://rgai4phalo.labeling.us-east-1.sagemaker.aws> [ ]

액 클라이언트  
1d3272645r7ej94hr18uaedque [ ]

**프라이빗 팀 (1) 정보**

프라이빗 작업 인력의 작업 팀입니다. 레이블 지정 작업을 위해 여러 프라이빗 팀을 만들 수 있습니다.

작업자	삭제	프라이빗 팀 만들기			
이름	▼	ARN	▼	생성 시간	▼
private-team-1	arn:aws:sagemaker:us-east-1:556463194331:workteam/private-crowd/private-team-1	2020년 7월 29일 오전 8:57 UTC			

**작업자 정보**

프라이빗 작업 인력의 모든 작업자

작업자	활성화	비활성화	삭제	새 작업자 초대					
작업자 검색	< 1 ... >	[ ]							
이메일	▲	상태	▼	Cognito 상태	▼	활성	▼	사용자 이름	▼
ggamini@hmail.hanyang.ac.kr		초대장 보냄		Force_change_password	예	ggamini@hmail.hanyang.ac.kr			

<https://console.aws.amazon.com/sagemaker/groundtruth?region=us-east-1#>

© 2008 - 2020, Amazon Web Services, Inc. 또는 계열사. All rights reserved. 개인 정보 보호 정책 이용 약관

- private team을 생성하면 다음과 같이 나옵니다.
- 메일로 들어가서 임시 password로 authorize를 해 줍니다.

The screenshot shows the AWS SageMaker Ground Truth interface. On the left, there's a sidebar with various options like 'Ground Truth' (selected), 'Notebook', 'Training', 'Tuning', 'Batch', and 'Jobs'. The main area displays 'Private Team' information, including the ARN of the Cognito user pool and a link to the AWS Lambda URL for logging in. Below this, it shows a 'Private Team (1)' table with one entry: 'private-team-1' (ARN: arn:aws:sagemaker:us-east-1:556463194331:workteam/private-crowd/private-team-1) created on 2020-07-29 08:57 UTC. Under the 'Workers' section, there's a table with one worker entry: 'ggamini@hmail.hanyang.ac.kr' (Cognito status: Confirmed). The bottom of the screen shows the URL https://console.aws.amazon.com/sagemaker/groundtruth?region=us-east-1# and some footer text.

- 위와 같이 authorized로 바뀝니다.

## 2. Create a labeling job

이제 labeling을 해보겠습니다.

The screenshot shows the AWS SageMaker console with the 'Labeling Jobs' section selected. The top navigation bar includes 'Amazon SageMaker Studio' and '레이블 지정 작업'. The main area has a search bar for 'Search labeling jobs' and a red button labeled '레이블 지정 작업 생성'. Below the search bar is a table header with columns: '이름' (Name), '상태' (Status), '작업 유형' (Job Type), '레이블이 지정된 객체/전체' (Labeled objects/whole), and '생성 시간' (Creation time). A message at the bottom states '현재 생성된 레이블 지정 작업이 없습니다.' (No labeling jobs have been created yet.). The bottom of the screen shows the URL https://console.aws.amazon.com/sagemaker/jobs?region=us-east-1# and some footer text.

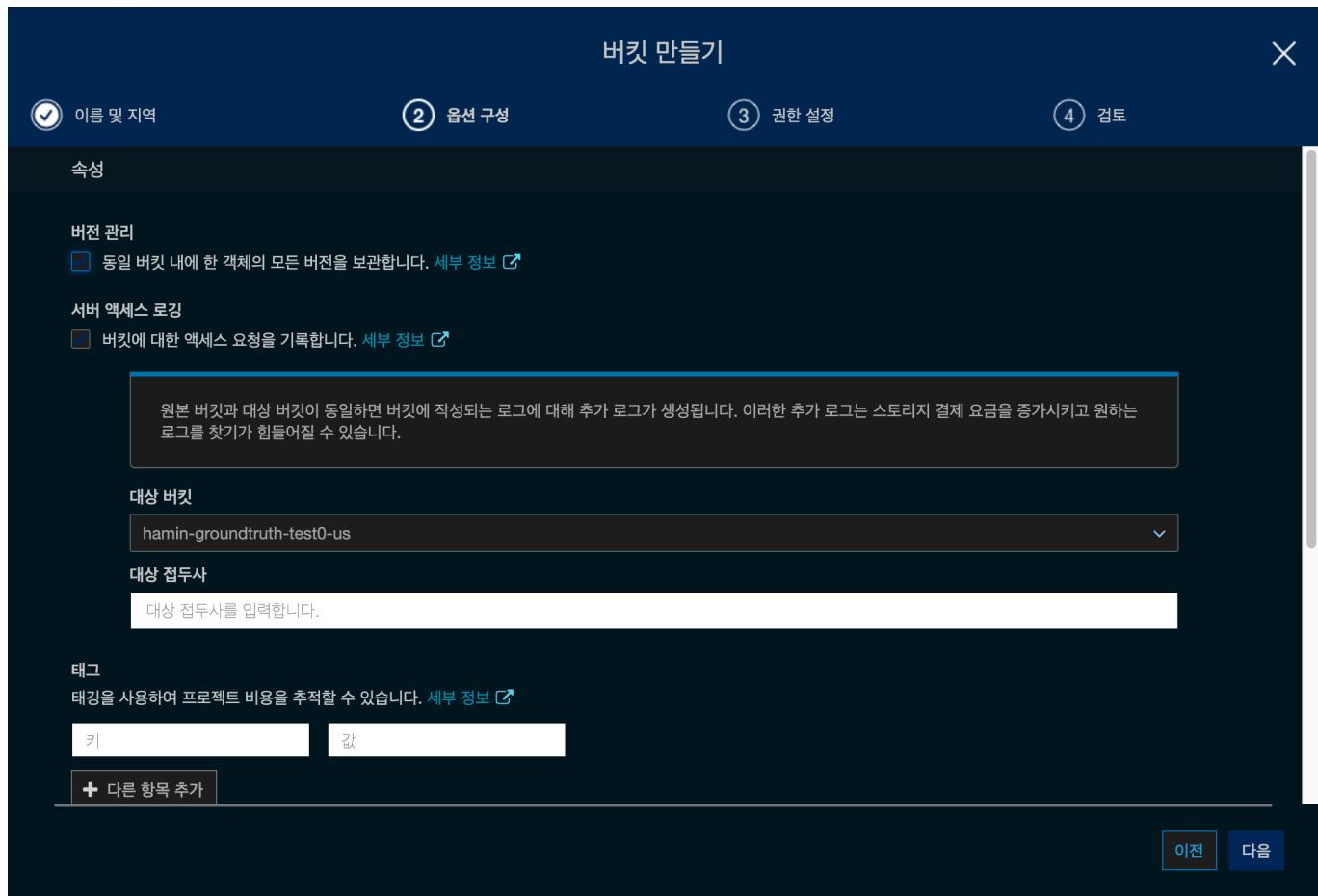
- 레이블 지정 작업 생성으로 들어갑니다.

The screenshot shows the '작업 세부 정보 지정' (Job Sub-Information Configuration) step in the SageMaker labeling job creation process. The left sidebar shows 'Step 1: 작업 세부 정보 지정' and 'Step 2: 작업자와 구성 도구 선택'. The main area is titled '작업 개요' (Job Overview). It includes fields for '작업 이름' (Job Name) set to 'job-demo-1', a note about character limits, and a checkbox for specifying a different label attribute name. Below this is the 'Input data setup' section, which offers 'Automated data setup' (selected) or 'Manual data setup'. The 'S3 location for input datasets' field contains 's3://input-dataset-bucket/optional-prefix'. The 'S3 location for output datasets' section shows 'Same location as input dataset' selected. The 'Data type' dropdown is set to 'Image'. At the bottom, there are links for ' 의견' (Feedback), ' 한국어' (Korean), and copyright information.

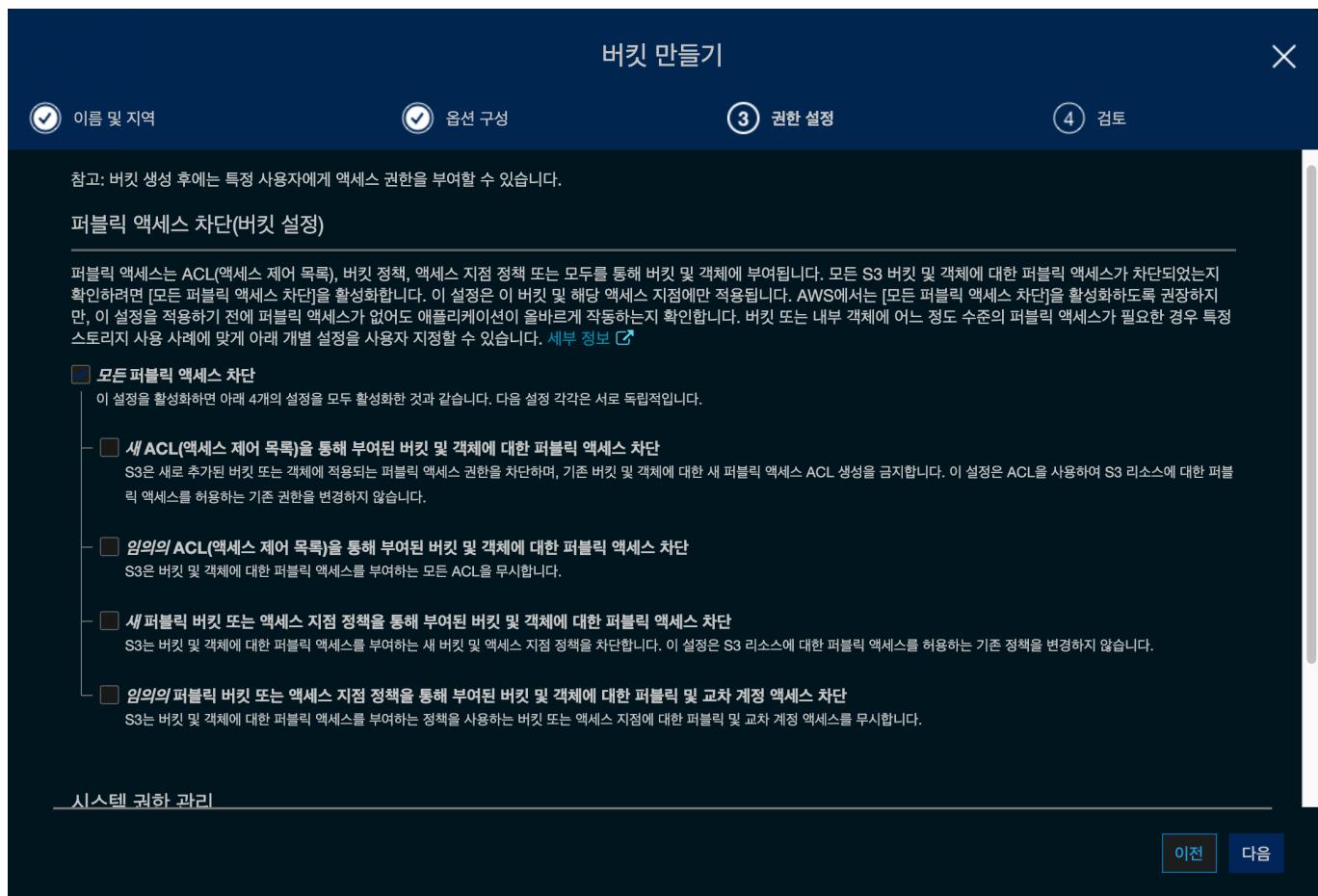
- s3에 있는 파일로 할 수도, manifest file을 만들 수도 있습니다.

The screenshot shows the '버킷 만들기' (Create Bucket) wizard, Step 1: 이름 및 지역 (Bucket Name and Region). The steps are numbered 1 through 4. The first step is completed, showing '버킷 이름' (Bucket Name) as 'hamin-groundtruth-test0-us' and '리전' (Region) as '미국 동부(버지니아 북부)'. The second step, '옵션 구성' (Option Configuration), is shown below. The third step, '권한 설정' (Permissions), and the fourth step, '검토' (Review), are also visible at the top. At the bottom are '생성' (Create) and '취소' (Cancel) buttons.

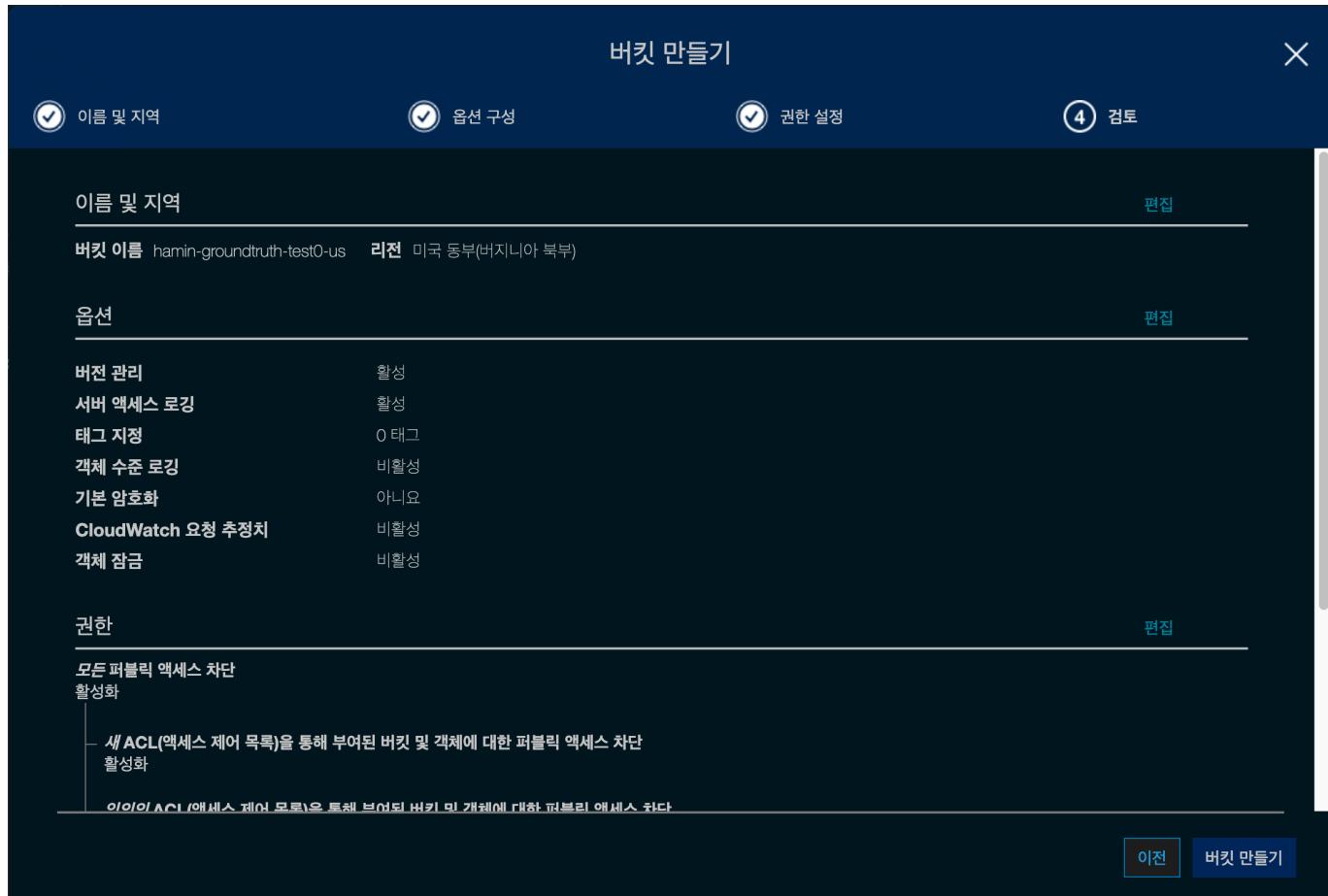
- S3는 다음과 같은 세팅을 하였습니다.



- 해당 버킷에 대한 액세스 요청을 모두 기록합니다.



- 권한 설정도 우선 기본 사항으로 하였습니다.



- 이렇게 S3 버킷을 생성했습니다.
- 그리고 샘플 dataset은 제 마음대로 넣어봤습니다.
- S3 버킷에 파일을 넣고 manifest를 생성해주어야 합니다.
- manifest 만들기는 다음을 참고합니다

[https://docs.aws.amazon.com/ko\\_kr/sagemaker/latest/dg/sms-data-input.html#sms-console-create-manifest-file](https://docs.aws.amazon.com/ko_kr/sagemaker/latest/dg/sms-data-input.html#sms-console-create-manifest-file)

<https://docs.aws.amazon.com/sagemaker/latest/dg/sms-getting-started-step1.html>

- manifest는 샘플들을 리스트하는 json 파일인데 docs와 다르게 manifest 없이도 할 수 있도록 되어있습니다.
- 하지만 저는 manifest를 직접 만들어 해보았습니다.

그리고 다시 labeling으로 돌아가

- 이런식으로 세팅을 해 줍니다

- dataset이 이미지를 이므로 task category는 image, task selection은 semantic segmentation으로 해보겠습니다.

다음으로 넘어가서

Amazon SageMaker Studio

Amazon SageMaker > 레이블 지정 작업 > 레이블 지정 작업 생성

작업자와 구성 도구 선택

작업자 유형

- 퍼블릭
- 프라이빗
- 공급업체 관리

프라이빗 팀

테스크 제한 시간

추가 구성 - 선택 사항

▶ 기존 레이블 표시 옵션 - 선택 사항

https://console.aws.amazon.com/sagemaker/home#/landing

- private으로 세팅해 줍니다.

의미 체계 분할 레이블 지정 도구

Good example

Enter description to explain a correctly done segmentation

Add image here

Bad example

Enter description of an incorrectly done segmentation

Add image here

▶ 추가 지침 - 선택 사항

▶ 추가 지침 - 선택 사항

취소 이전 생성

© 2008 – 2020, Amazon Web Services, Inc. 또는 계열사. All rights reserved. 개인 정보 보호 정책 이용 약관

- 이런식으로 예시가 나옵니다
- 구분해야할 물체를 최대한 구체적으로 명시해줍니다.

Amazon SageMaker Studio

Amazon SageMaker > 레이블 지정 작업

작업	레이블 지정 작업
<a href="#">C</a>	작업
<a href="#">Search labeling jobs</a>	레이블 지정 작업 생성
이름 상태 작업 유형 job-demo-1 진행 중 의미 체계 분할 2020년 7월 30일 오전 9:19 UTC	

© 2008 - 2020, Amazon Web Services, Inc. 또는 계열사. All rights reserved. 개인 정보 보호 정책 이용 약관

- 생성 버튼을 눌러주면 다음과 같이 labeling job이 생성 됩니다.

Amazon SageMaker Studio

Amazon SageMaker > 레이블 지정된 데이터 세트 객체 (-)

출력 | 태그

작업	레이블 지정된 데이터 세트 객체 (-)
<a href="#">pic1.jpg</a>	
<a href="#">pic2.jpg</a>	
<a href="#">pic3.jpg</a>	
<a href="#">pic4.jpg</a>	

© 2008 - 2020, Amazon Web Services, Inc. 또는 계열사. All rights reserved. 개인 정보 보호 정책 이용 약관

- 들어가보면 작업 중인 사진들이 보입니다.
- 시간이 좀 걸리니 기다려야 합니다.

Amazon SageMaker > 레이블 지정 작업

작업	레이블 지정 작업 생성
job-demo-1	2020년 7월 30일 오전 9:19 UTC

- 결과물을 보려면 console로 들어가야 합니다.

Amazon SageMaker > 레이블 지정 작업 인력

작업자	생성 시간
private-team-	2020년 7월 29일 오전 8:57 UTC

- 오른쪽의 레이블 지정 포털 로그인 URL로 들어가면 됩니다.

Amazon SageMaker Ground Truth - Worker Portal

Hello, ggaminig@mail.hanyang.ac.kr

Log out Show instructions Start working

Jobs

Description	Customer ID	Creation time
Semantic segmentation: please segment people and traffic light and bus, tree, car	556463194331	July 30, 2020 9:21:21 UTC

- 들어가면 다음과 같은 화면이 나옵니다.
- start working을 들어가면

Amazon SageMaker Ground Truth - Worker Portal

Hello, ggaminig@mail.hanyang.ac.kr Customer ID: 5564... Task description: Draw pixel level labels around specific objects and... Task time: 0:42 of 60 Min Stop working Log out

Instructions: please segment people and traffic light and bus, tree, car

**Instructions**

[View full instructions](#)  
[View tool guide](#)  
[How to use the Auto-segment tool](#)

**Good example**

Enter description to explain a correctly done segmentation

Add image here

**Bad example**

Enter description of an incorrectly done segmentation

Add image here

**Labels**

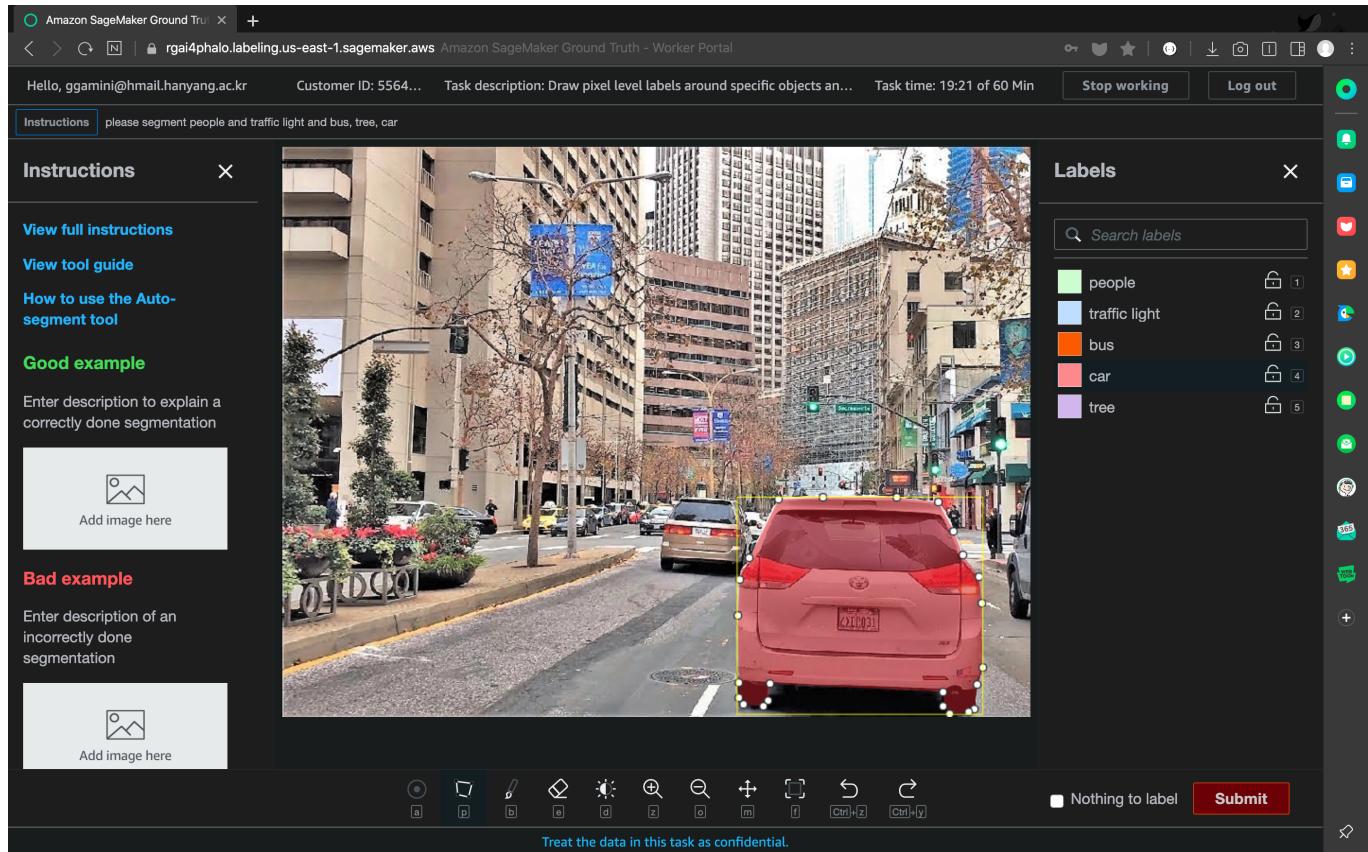
Search labels

- people
- traffic light
- bus
- car
- tree

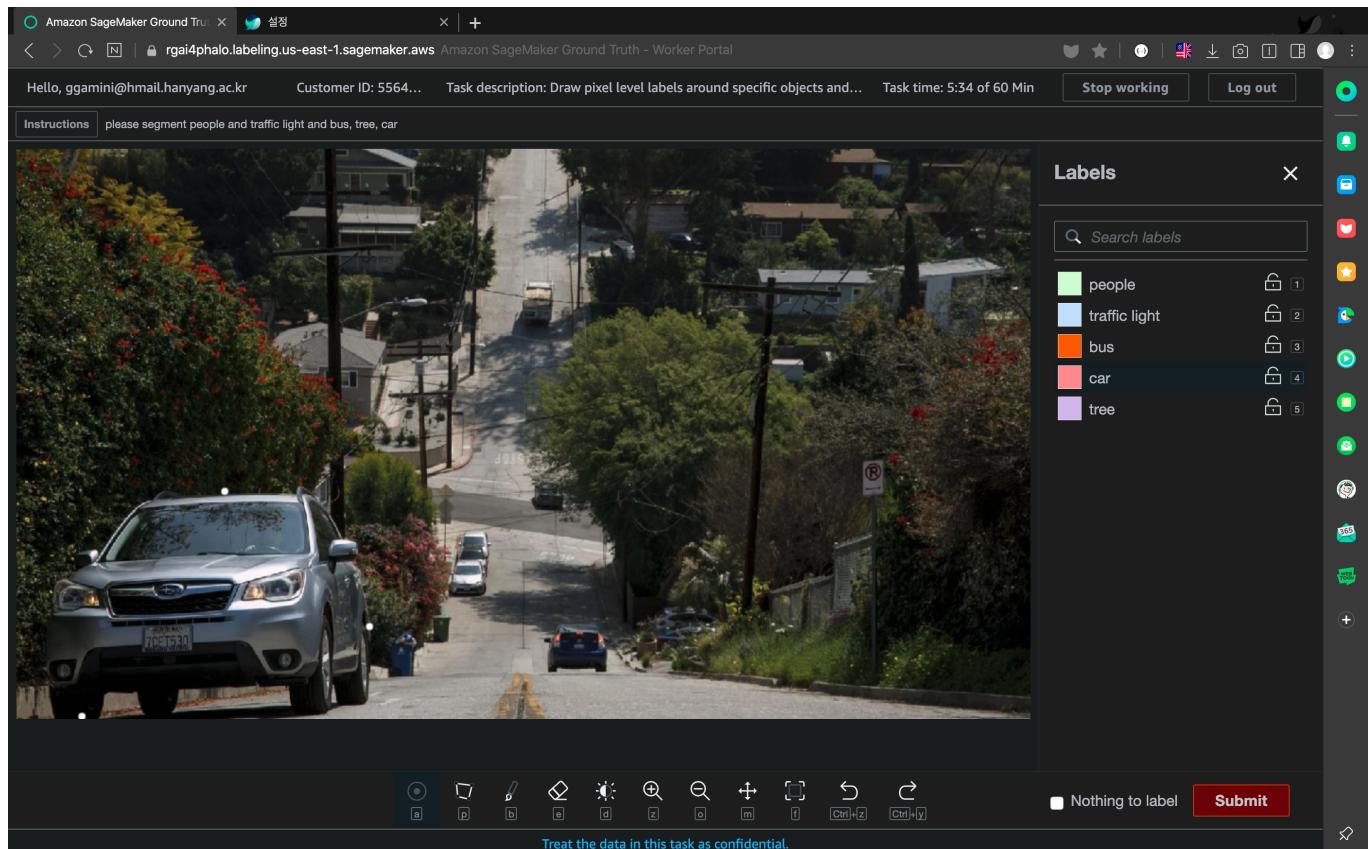
Treat the data in this task as confidential.

Submit

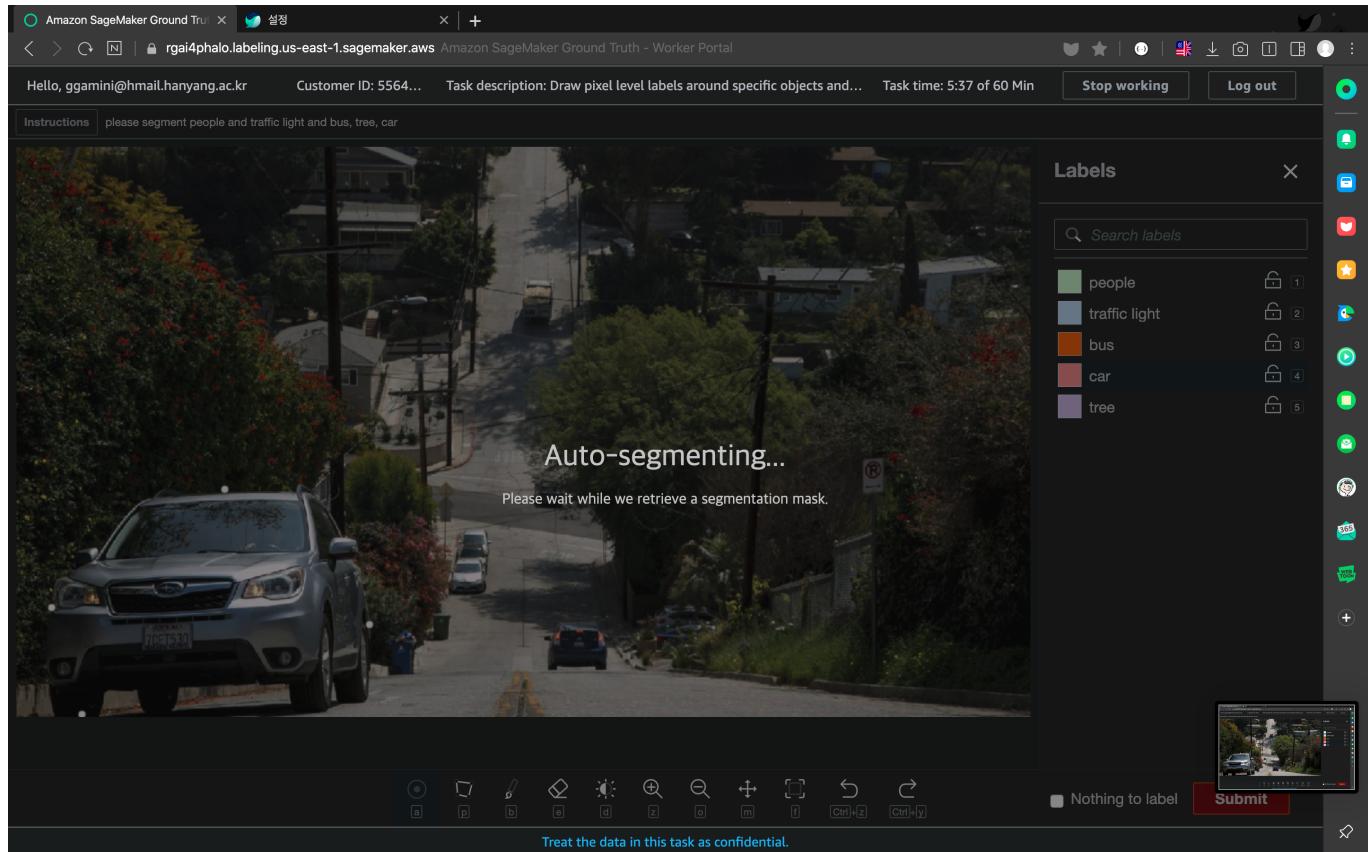
- 다음과 같은 화면이 나옵니다.



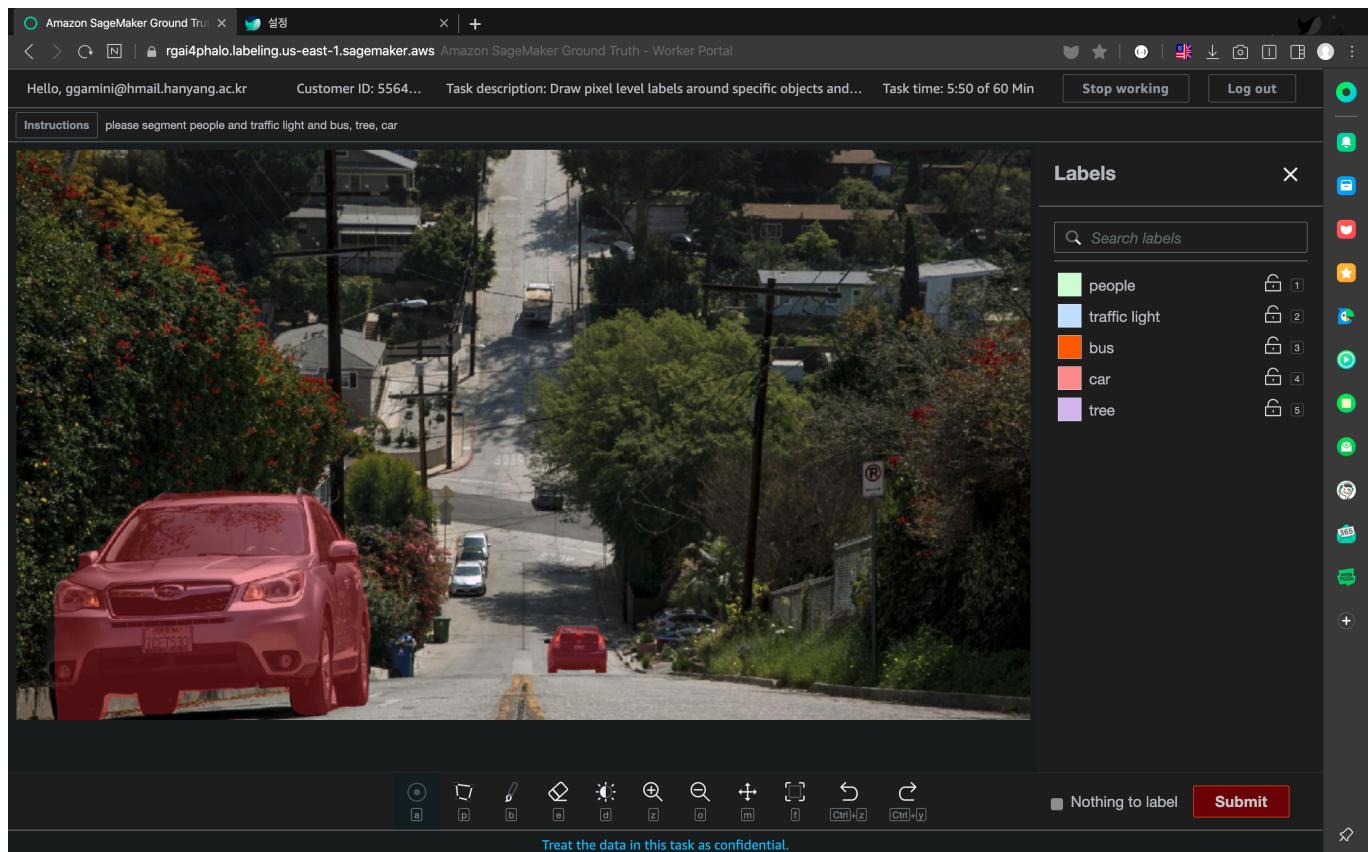
- 이런식으로 지정해줍니다
- 그리고 submit을 해줍니다.



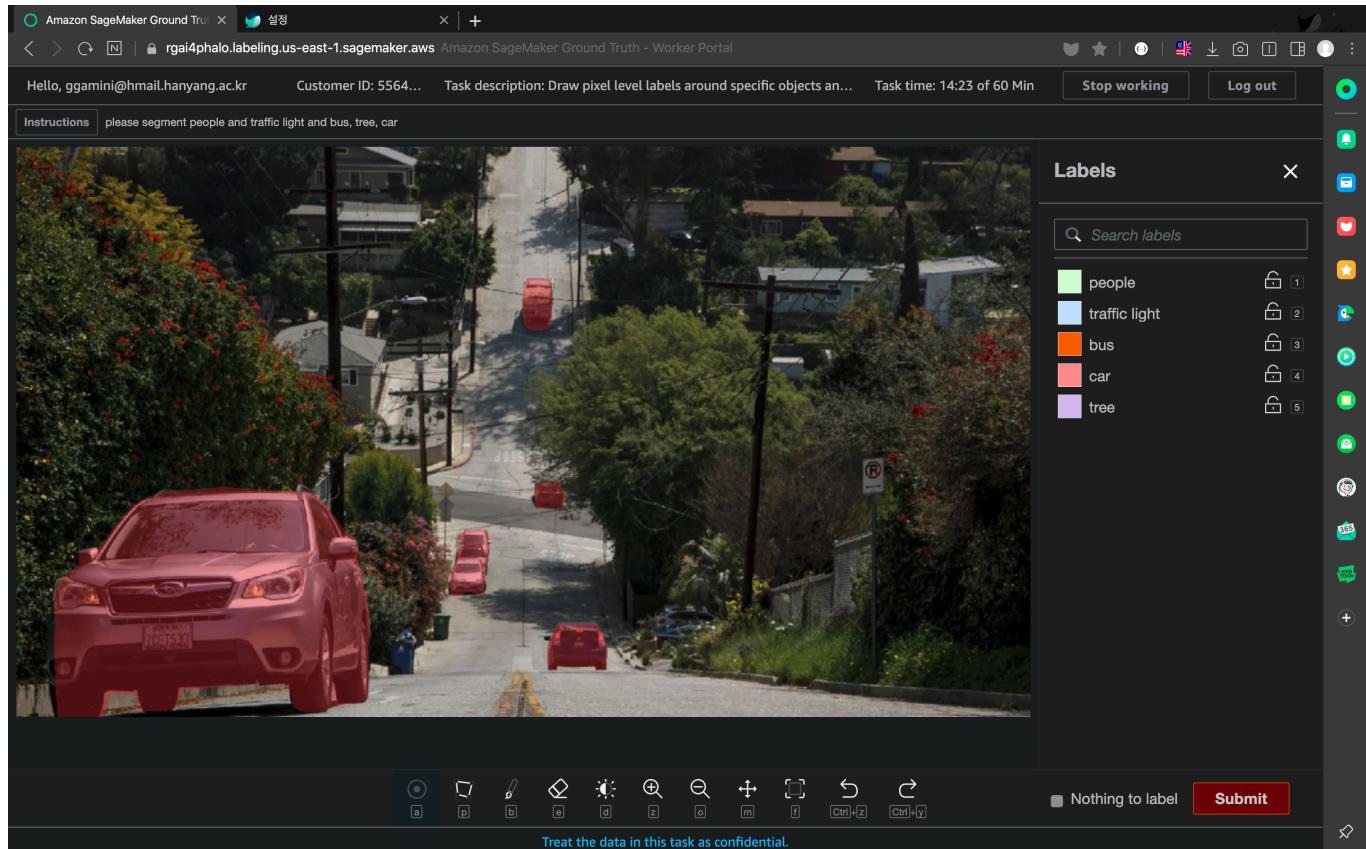
- 이번엔 auto segment 기능을 이용해보겠습니다.



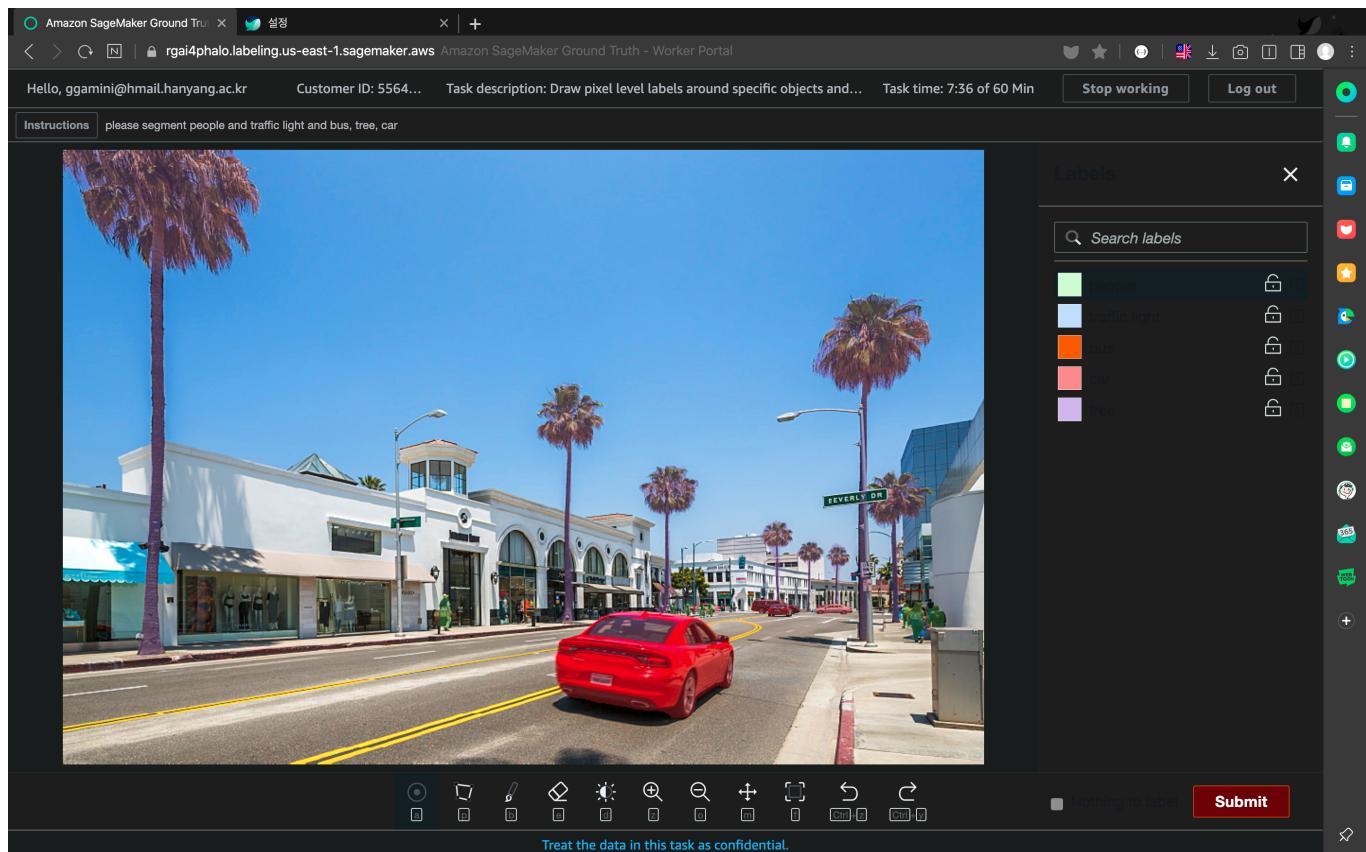
- Labels에서 해당되는것을 클릭후 최하단, 최상단, 좌우 끝 네 지점을 지정해주면 auto segmenting이 실행됩니다.



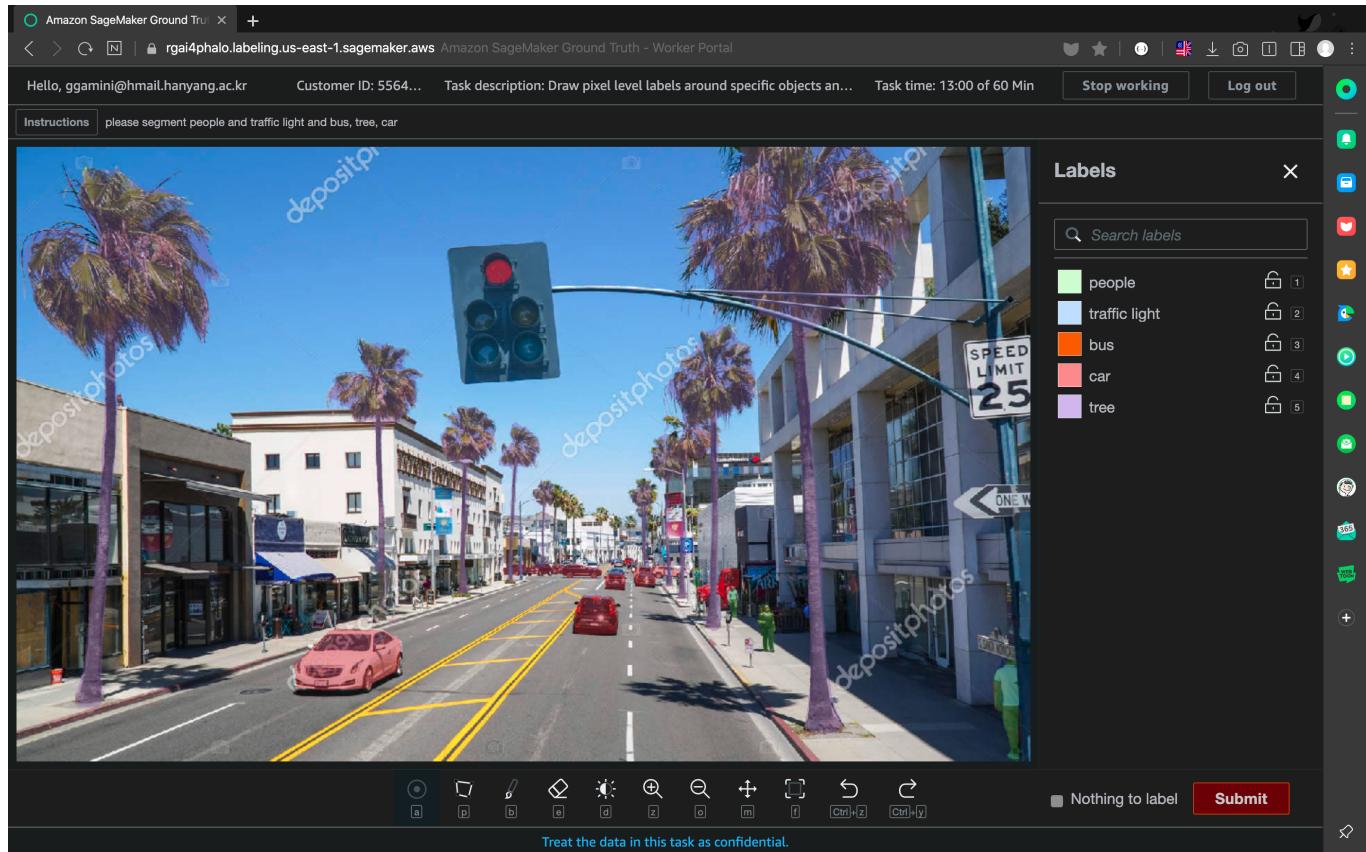
- 이런식으로 잘 지정이 됩니다.



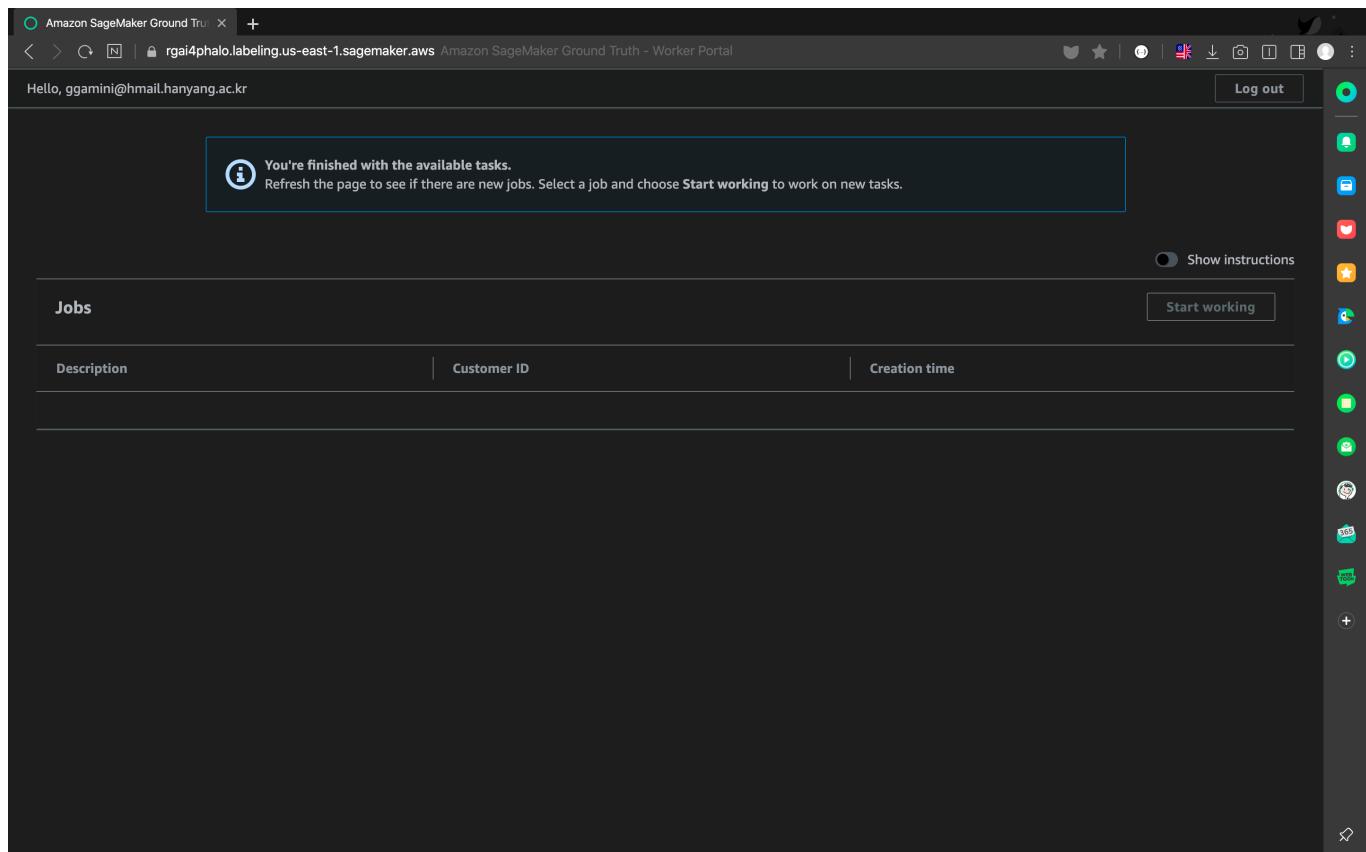
- 다른 차들도 해줍니다.



- 다른 요소들도 해줍니다.



- 다 해줍니다.



- 모두 마무리하면 이런 화면으로 넘어갑니다.

The screenshot shows the AWS CloudWatch Management Console interface. On the left, there's a sidebar with various AWS services like CloudWatch, ServiceLens, Container Insights, and Synthetics. The main area is titled 'Log events' and shows a list of log entries for a job named 'job-demo-1'. The logs are timestamped from August 2020, showing progress messages such as 'Labeling job progress for batch number:1 | #Total data set objects:8 | #Data'. A search bar at the top allows filtering by event name.

- cloud watch에 들어가면 실행한 log를 볼 수 있습니다.

The screenshot shows the AWS S3 Management Console interface. It displays the contents of a bucket named 'hamin-groundtruth-test0-us'. Under the 'output' folder, several files are listed, including 'index', 'input', and several log files starting with '2020-08-02-07-40-'. The table below provides a detailed view of these files:

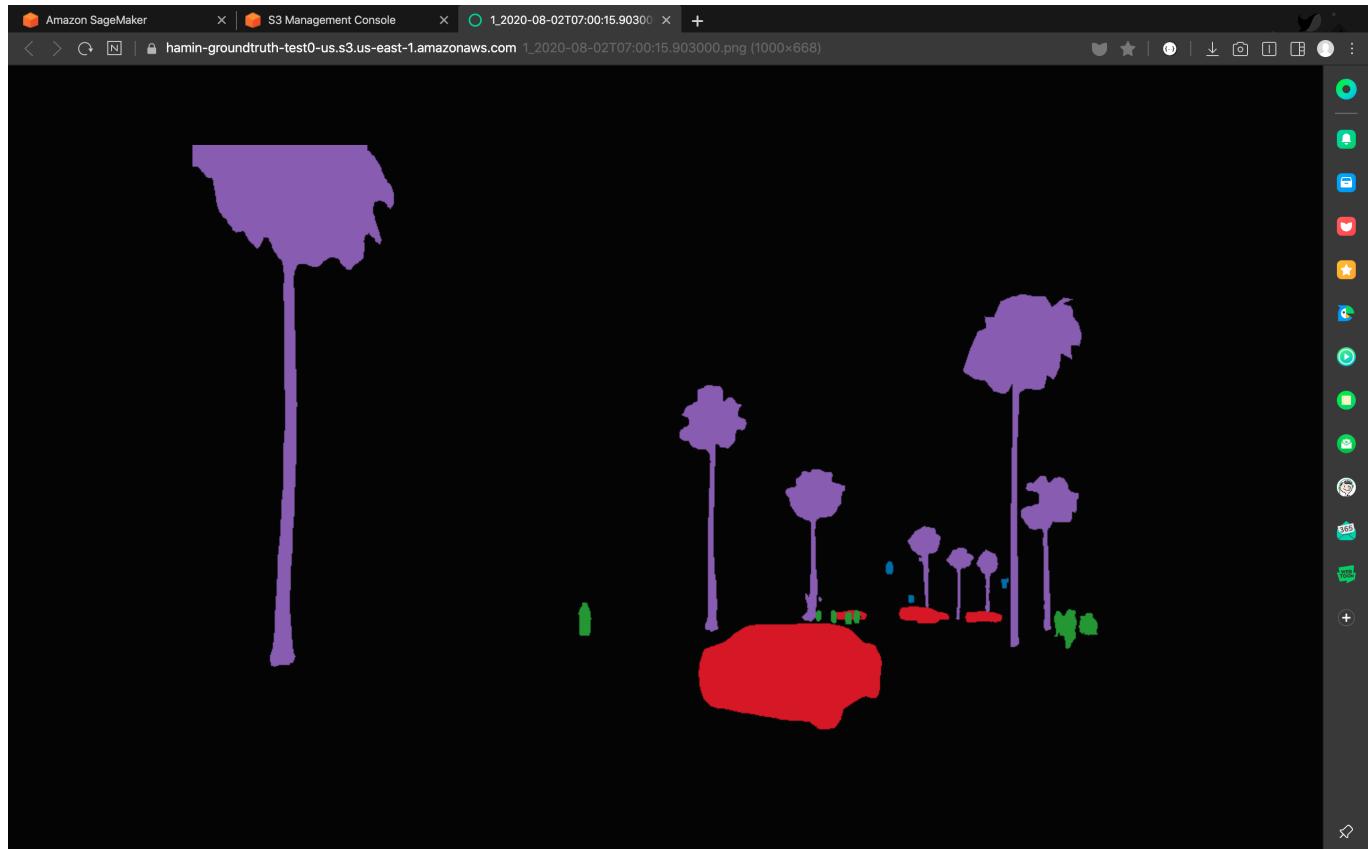
이름	마지막 수정	크기	스토리지 클래스
index	--	--	--
input	--	--	--
output	--	--	--
2020-08-02-07-40-06-A11479F487ADA293	8월 2, 2020 4:40:07 오후 GMT+0900	719.0 B	스탠다드
2020-08-02-07-40-08-64CD4E54E7A1970A	8월 2, 2020 4:40:09 오후 GMT+0900	1.7 KB	스탠다드
2020-08-02-07-40-21-1D03178B4740E18E	8월 2, 2020 4:40:22 오후 GMT+0900	1.7 KB	스탠다드
2020-08-02-07-40-23-D3037BE3268570EB	8월 2, 2020 4:40:24 오후 GMT+0900	685.0 B	스탠다드
2020-08-02-07-40-26-743026E331AD20E2	8월 2, 2020 4:40:27 오후 GMT+0900	1.7 KB	스탠다드
2020-08-02-07-40-31-5D3B19CF3CF064D1	8월 2, 2020 4:40:32 오후 GMT+0900	1.7 KB	스탠다드

- S3의 output에 들어가면 결과물이 들어있습니다.

The screenshot shows the AWS S3 Management Console interface. The path in the navigation bar is: Amazon S3 > hamin-groundtruth-test0-us > output > job-demo-1 > annotations > consolidated-annotation > output. The main area displays a list of annotated images. The columns include Name, Last modified, Size, and Storage Class. One file, '1\_2020-08-02T07:00:15.903000.png', is selected and highlighted in blue.

이름	마지막 수정	크기	스토리지 클래스
0_2020-08-02T07:38:26.700000.png	8월 2, 2020 4:38:27 오후 GMT+0900	12.8 KB	стандарт
1_2020-08-02T07:00:15.903000.png	8월 2, 2020 4:00:16 오후 GMT+0900	4.8 KB	стандарт
2_2020-08-02T07:14:42.914000.png	8월 2, 2020 4:14:43 오후 GMT+0900	11.3 KB	стандарт
3_2020-08-02T06:30:15.279000.png	8월 2, 2020 3:30:16 오후 GMT+0900	937.0 B	стандарт
4_2020-08-02T07:31:14.293000.png	8월 2, 2020 4:31:15 오후 GMT+0900	31.6 KB	стандарт
5_2020-08-02T06:45:46.945000.png	8월 2, 2020 3:45:47 오후 GMT+0900	2.0 KB	стандарт
6_2020-08-02T07:17:49.378000.png	8월 2, 2020 4:17:50 오후 GMT+0900	3.4 KB	стандарт
7_2020-08-02T06:51:59.428000.png	8월 2, 2020 3:52:00 오후 GMT+0900	3.0 KB	стандарт

- 결과물들은 이곳에 있습니다.



- 이런식으로 있습니다.

The screenshot shows the Amazon SageMaker Studio interface. On the left, a sidebar lists various project categories: Ground Truth (Labeling Job, Dataset with Labeling, Labeling Inference), Notebooks (Notebook Instances, Configuration, Git Repository), Processing (Job Processing, Model Packaging, Model), and Tuning (Algorithm Tuning, Hyperparameter Tuning). The main content area displays a 'job-demo-1' job configuration. It includes sections for 'Labeling Job Details' (Input Dataset Location: s3://hamin-groundtruth-test0-us/input/index/dataset.manifest.json, Output Dataset Location: s3://hamin-groundtruth-test0-us/output/job-demo-1/), 'ARN' (arn:aws:sagemaker:us-east-1:55646319431:labeling-job/job-demo-1), and 'Job Type' (Private). Below this, a 'Labeling Data Set Objects (8)' section shows four preview images: a street scene with green labels, a red car, a street scene with palm trees, and a red car from a different angle. A legend at the bottom right indicates the labels used.

- 이곳에 가면 output의 경로를 볼 수 있습니다.

The screenshot shows the Amazon SageMaker Studio interface. The left sidebar is identical to the previous one. The main content area displays a 'Labeling Data Set Objects (8)' section. It shows four images with labels applied: 'pic1.jpg' (a street scene with green labels), 'pic2.jpg' (a red car), 'pic3.jpg' (a street scene with palm trees and a watermark for depositphotos), and 'pic4.jpg' (a red car from a different angle). The interface includes navigation arrows and a legend at the bottom right.

- 이런식으로 잘 labeling이 되어있습니다.

결과물을 보면



i| 사진의 경우



```
1  {"answers": [{"answerContent":  
    "crowd-semantic-segmentation":  
    "inputImageProperties": {"height": 774,  
    "width": 1160}, "labelMappings": {"bus":  
    {"color": "#ff7f0e"}, "car": {"color": "#d62728"},  
    "people": {"color": "#2ca02c"}, "traffic light":  
    {"color": "#1f77b4"}, "tree": {"color": "#9467bd"}},  
    "labeledImage":  
    {"pngImageData": "iVBORw0KGgoAAAANSUhEUgAABAgAAAM  
    GBAMAAABrH6LWAAAAAXNSR0IB2cksfwAAABJQTFRFAAAA1Ge  
    9H3e0/380LKAs1icop8xtTgAAAAZ0Uk5TAP//enng/  
    gAAMq5JREFUeJztnUt2I7m0Q0VM9tzeQXauoM7JWgDfafW8J  
    7X/  
    rXRJVij4AQiABH8S7qRKChIAwZsRIVmWL5cvXT4vxvthEhnN  
    mERGMyaR0YyyRF+z120MR9shk  
    +gNMYmMZkwioxmTyGj GJDKaMYmMZtQlsjeK3g  
    +TyGj GJDKaMYmMZkwioxmTyGj GJDKaMYmMZkwioxmTyGj GJD  
    KaMYmMZkwioxmTyGhGXSL7LMj 7YRIZzehLZNezt8MkMppZRK  
    IvuwxuzAIHRN7LM8YwXSJgpl9Vmh0Z7ZEDfoZqzBZonhqr0  
    UafZkr0YdJ9AroSyRQIXHIrmeb0kEitgqpQybRpsyRCJvaf7  
    1GB2ZILJ2BTKK96SARqUL9TGNJJkiEn4hMoj2ZIFFhpt1zb4  
    LJZDTTQ6KyCoWrmlUm0J2tJZDdFWzJeouLUQas2VF1MItrue7U  
    gPiYrnk  
    +LVzCTaki4SFVQo02QWbcloici545ZuaNFFItwi6kRkEu3Ic  
    hLZ9Ww/BkvUMNVYlrES0Sciu55tyFiJOHOHLt/  
    QoI9EiAmsqXY9246RErXMNRamk0TQ6YRzQ2QS7UgniSATTKJ  
    XZZxEXIdMou3oJVF  
    +PWuYaqyNSWQ000ui7KLEn2kS7cYoidh3RPtUY3VGsdQw1Vi  
    dXg61SGTXs83o5VAikeRqZhLtRi  
    +HvhrlSHBJ9mE170Emh9HQim/  
    0ECDLJLJLIC5C1D1QHBDWLGEXGCCGXYBDD1O.../2
```

Ln 1, Col 1   Spaces: 4   UTF-8   LF   JSON   

이런식으로 라벨이 붙어있습니다.

## 이미지 분류 (다중 레이블)

- 이번엔 이미지 분류 다중 레이블로 해보겠습니다

위에 했던 방식과 똑같이 해주면

Description	Customer ID	Creation time
● Image Classification (Multi-label): please segment people , car, bus, tree, traffic light, traffic sign	556463194331	August 3, 2020 8:16:11 UTC

이렇게 워크 포털이 나옵니다 들어가서

AWS Account | Workbench | S3 Management | 이미지 분류(단일 라이브러리) | bithumb\_bot/3\_ | Amazon SageMaker | Amazon SageMaker | (112) aws sagemaker | + | Twitter | Star | Help | Download | Print | Copy | Refresh | Log out

Hello, ggamini@hmail.hanyang.ac.kr Customer ID: 556463194331 Task description: Categorize images into multiple classes Task time: 0:13 of 5 Min Stop working Log out

Instructions: please segment people, car, bus, tree, traffic light, traffic sign



Select appropriate categories

people	1
car	2
bus	3
tree	4
traffic light	5
traffic sign	6

Submit

Treat the data in this task as confidential.

지정해 주면 됩니다

AWS Account | Workbench | S3 Management | 이미지 분류(단일 라이브러리) | bithumb\_bot/3\_ | Amazon SageMaker | Amazon SageMaker | (112) aws sagemaker | + | Twitter | Star | Help | Download | Print | Copy | Refresh | Log out

Hello, ggamini@hmail.hanyang.ac.kr Customer ID: 556463194331 Task description: Categorize images into multiple classes Task time: 0:15 of 5 Min Stop working Log out

Instructions: please segment people, car, bus, tree, traffic light, traffic sign



Select appropriate categories

people	1
car	2
bus	3
tree	4
traffic light	5
traffic sign	6

Submit

Treat the data in this task as confidential.

계속 지정해 주면 됩니다

AWS Account | Workbench | S3 Management | 이미지 분류(단일화) | bitumb\_bot/3.ti | Amazon SageM... | Amazon SageM... | (112) aws sagem... | + | rgai4phalo.labeling.us-east-1.sagemaker.aws Amazon SageMaker Ground Truth - Worker Portal

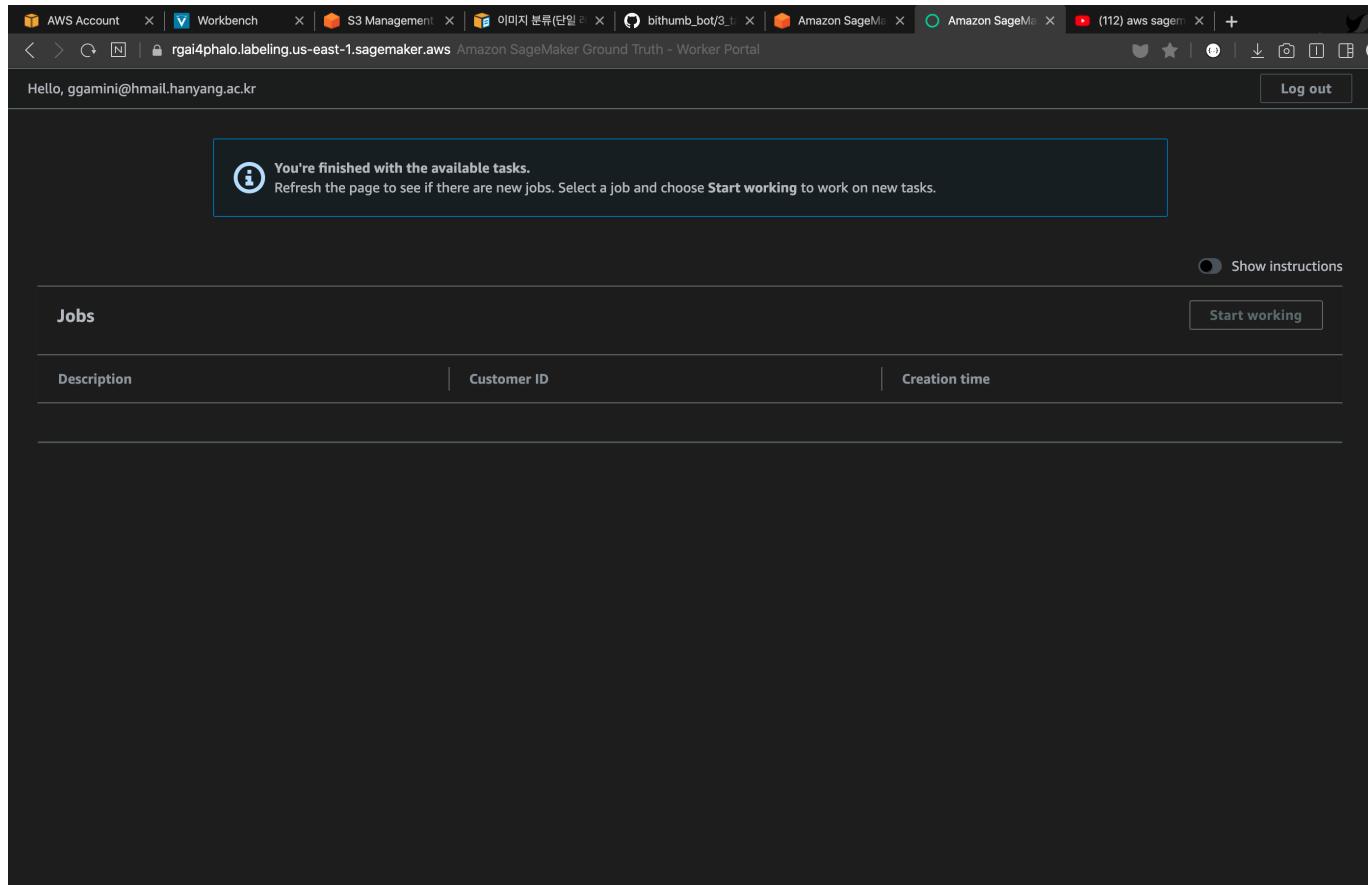
Hello, ggamini@hmail.hanyang.ac.kr

You're finished with the available tasks.  
Refresh the page to see if there are new jobs. Select a job and choose Start working to work on new tasks.

Show instructions | Log out

**Jobs**

Description | Customer ID | Creation time



작업이 끝났습니다

결과물을 보면



이 사진의 경우

```
AWS > groundTruth > {} 2020-08-08_08_27_59.json > ...
1   ":{"labels":["people","car","bus","tree","traffic light"]}},"submissionTime":"2020-08-03T08:27:40.664Z","work...
```

이런식으로 사람, 자동차, 버스, 나무, 신호등의 label이 붙습니다.

이러한 라벨들을 다른곳에 활용할 수 있을 것입니다.