

Object Detection용 주요 Dataset

1. PASCAL VOC Dataset

주요 Dataset 소개

많은 Detection과 Segmentation 딥러닝 패키지가 아래 Dataset들을 기반으로 Pretrained 되어 배포

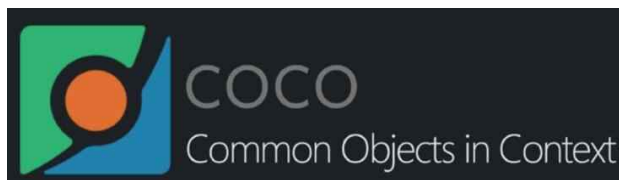
PASCAL VOC



XML Format

20개의 오브젝트 카테고리

MS COCO



json Format 80개

의 오브젝트 카테고리

Google Open Images



csv Format

600개의 오브젝트 카테고리

Pascal VOC Dataset

PASCAL VOC는 2005년에서 2012년까지 진행되었던 PASCAL VOC Challenge에서 쓰이던 데이터셋이다 그 중 PASCAL 2007과 PASCAL 2012 데이터셋이 벤치마크 데이터셋으로 자주 사용됨.

- classification Dataset: 클래스로 이름으로 된 폴더별로 이미지
- localization Dataset: 이미지 한 장에 좌표 값과 클래스

1. Annotation

- XML 포맷 :
 - Image info : width, height, depth, name, pose
 - Bounding box 좌표값 4개 xmin,ymin,xmax,ymax

PACAL VOC 데이터셋 크기

Annotation , ImageSet , JPEGImages

- 이미지 개수:
 - object detection: 총 9,963개의 주석이 달린(annotated) 이미지가 포함. 이 중에서 5,011개가 학습 데이터
 - segmentation: 422개의 학습 데이터
- 이미지당 평균 object 수: 2.4개
- 이미지당 평균 class 수: 1.4개
- class 개수 20개

PASCAL VOC 2012

<http://host.robots.ox.ac.uk/pascal/VOC/voc2012/>



Visual Object Classes Challenge 2012 (VOC2012)

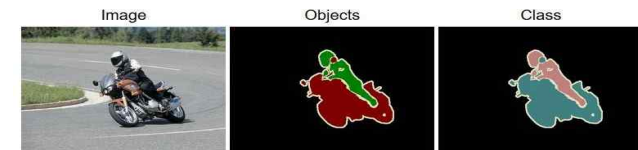


[click on an image to see the annotation]

Classification
/Detection



Segmentation



Action Cl
assification



Person Layout

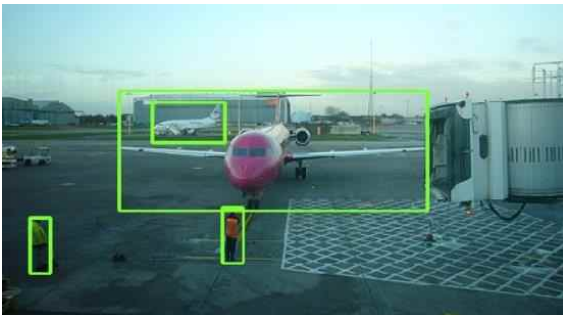


Annotation 이란?

이미지의 Detection 정보를 별도의 설명 파일로 제공되는 것을 일반적으로 Annotation 이라고 함.

Annotation은 Object의 Bounding Box 위치나 Object 이름등을 특정 포맷으로 제공함.

원본 이미지에 Bounding Box 시각화



원본 이미지



Annotation

```
<annotation>
  <folder>VOC2012</folder>
  <filename>2007_000032.jpg</filename>
  <source>
    <database>The VOC2007 Database</database>
    <annotation>PASCAL VOC2007</annotation>
    <image>flickr</image>
  </source>
  <size>
    <width>500</width>
    <height>281</height>
    <depth>3</depth>
  </size>
  <segmented>1</segmented>
  <object>
    <name>aeroplane</name>
    <pose>Frontal</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>104</xmin>
      <ymin>78</ymin>
      <xmax>375</xmax>
      <ymax>183</ymax>
    </bndbox>
  </object>
  <object>
    <name>aeroplane</name>
    <pose>Left</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>133</xmin>
      <ymin>88</ymin>
      <xmax>197</xmax>
      <ymax>123</ymax>
    </bndbox>
  </object>
</annotation>
```

PASCAL VOC Dataset 구조

VOC 2012 기준

VOCdevkit

VOC2012

Annotations

Xml 포맷이며, 개별 xml파일은 한 개 image에 대한 Annotation 정보를 가지고 있음. 확장자 xml을 제외한 파일명은 image 파일명(확장자 jpg를 제외한)과 동일하게 매핑

ImageSet

어떤 이미지를 train, test, trainval, val에 사용할 것인지에 대한 매핑 정보를 개별 오브젝트별로 파일로 가지고 있음.

JPEGImages

Detection과 Segmentation에 사용될 원본 이미지

SegmentationClass

Semantic Segmentation에 사용될 masking 이미지

SegmentationObject

Instance Segmentation에 사용될 masking 이미지

Annotation 파일 예시

Annotation 파일 2007_000032.xml 파일 일부



2007_000032.jpg 파일에 대한 Annotation 정보를 가지고 있음.

```
<annotation>
  <folder>VOC2012</folder>
  <filename>2007_000032.jpg</filename>
  <source>
    <database>The VOC2007 Database</database>
    <annotation>PASCAL VOC2007</annotation>
    <image>flickr</image>
  </source>
  <size>
    <width>500</width>
    <height>281</height>
    <depth>3</depth>
  </size>
  <segmented>1</segmented>
  <object> ← 개별 Object 정보
    <name>aeroplane</name>
    <pose>Frontal</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>104</xmin>
      <ymin>78</ymin>
      <xmax>375</xmax>
      <ymax>183</ymax>
    </bndbox>
  </object>
  <object> ← 개별 Object 정보
    <name>aeroplane</name>
    <pose>Left</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>133</xmin>
      <ymin>88</ymin>
      <xmax>197</xmax>
      <ymax>123</ymax>
    </bndbox>
  </object>
</annotation>
```

개별 Object의
bounding box 정보 →

← 개별 Object 정보

개별 Object의
bounding box 정보 →

2. MS COCO Dataset

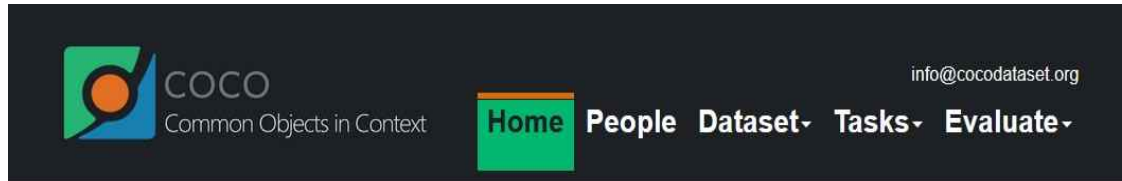
COCO Dataset

PASCAL VOC나 ImageNet이 현실세계 사진에서는 object를 잘 포착하지 못한다는 문제점이 있었다. 이런 문제점을 해결하기 위해 2014년 COCO 데이터셋이 등장한다

coco 데이터셋 구조

- JSON format
- Annotation:
 - Class
 - Bounding box (x,y,w,h)
 - Segmentation
- JPEG Images

MS-COCO Dataset 소개



<https://cocodataset.org/#home>

News

- We are pleased to announce the COCO 2019 [Detection](#), [Keypoint](#), [Panoptic](#), and [DensePose](#) Challenges.
- The new rules and awards for this year challenges encourage innovative methods.
- Results to be announced at the [Joint COCO and Mapillary Recognition ICCV workshop](#).

What is COCO?



COCO is a large-scale object detection, segmentation, and captioning dataset. COCO has several features:

- ✓ Object segmentation
- ✓ Recognition in context
- ✓ Superpixel stuff segmentation
- ✓ 330K images (>200K labeled)
- ✓ 1.5 million object instances
- ✓ 80 object categories
- ✓ 91 stuff categories
- ✓ 5 captions per image
- ✓ 250,000 people with keypoints

Collaborators

Tsung-Yi Lin Google Brain
Genevieve Patterson MSR, Trash TV
Matteo R. Ronchi Caltech
Yin Cui Google
Michael Maire TTI-Chicago
Serge Belongie Cornell Tech
Lubomir Bourdev WaveOne, Inc.
Ross Girshick FAIR
James Hays Georgia Tech
Pietro Perona Caltech
Deva Ramanan CMU
Larry Zitnick FAIR
Piotr Dollár FAIR

Sponsors



- 80개 Object Category
- 총 330K의 Image들과 1.5 Million 개의 object들
- (하나의 image에 평균 5개의 Object들로 구성)
- Tensorflow Object Detection API 및 많은 오픈 소스 계열의 주요 패키지들은 COCO Dataset으로 Pretrained 된 모델을 제공함

coco 데이터셋의 장점

- 다양한 크기의 물체가 존재
- 높은 비율로 작은 물체들이 존재
- Object들이 혼잡하게 존재하고, occlusion(폐색)이 많이 존재
- 덜 iconic 합니다. 여기서 iconic의 의미는 이미지가 특정 카테고리에 명확하게 속해 있는 경우를 말합니다. 예를 들어 아래 이미지의 (a)의 이미지들은 누가 봐도 명확하게 개, 소 등의 명확하게 특정 카테고리에 속한 걸 누구나 알 수 있습니다. (b)의 이미지들은 (a) 보다는 모호하지만, 특정 카테고리에 속하다고 말할 수 있습니다. 그런데 (c)의 경우 어떤 카테고리에 속하는지 분류하기가 모호합니다. 현실세계 사진에서는 (a)나 (b)보단 (c)에 속하는 경우가 훨씬 많이 존재합니다.



(a) Iconic object images

(b) Iconic scene images

(c) Non-iconic images

COCO Dataset 크기

- 이미지 개수:
 - 학습(training) 데이터셋: 118,000장의 이미지
 - 검증(validation) 데이터셋: 5,000장의 이미지
 - 테스트(test) 데이터셋: 41,000장의 이미지
- 이미지당 평균 object 수: 8개
- 이미지당 평균 class 수: 3.5개
- class 개수 91개(실제 사용은 80개)
- 수행하는 task
 - classification
 - object detection
 - semantic segmentation
 - instance segmentation
 - pose estimation etc

MS-COCO Dataset 오브젝트 카테고리

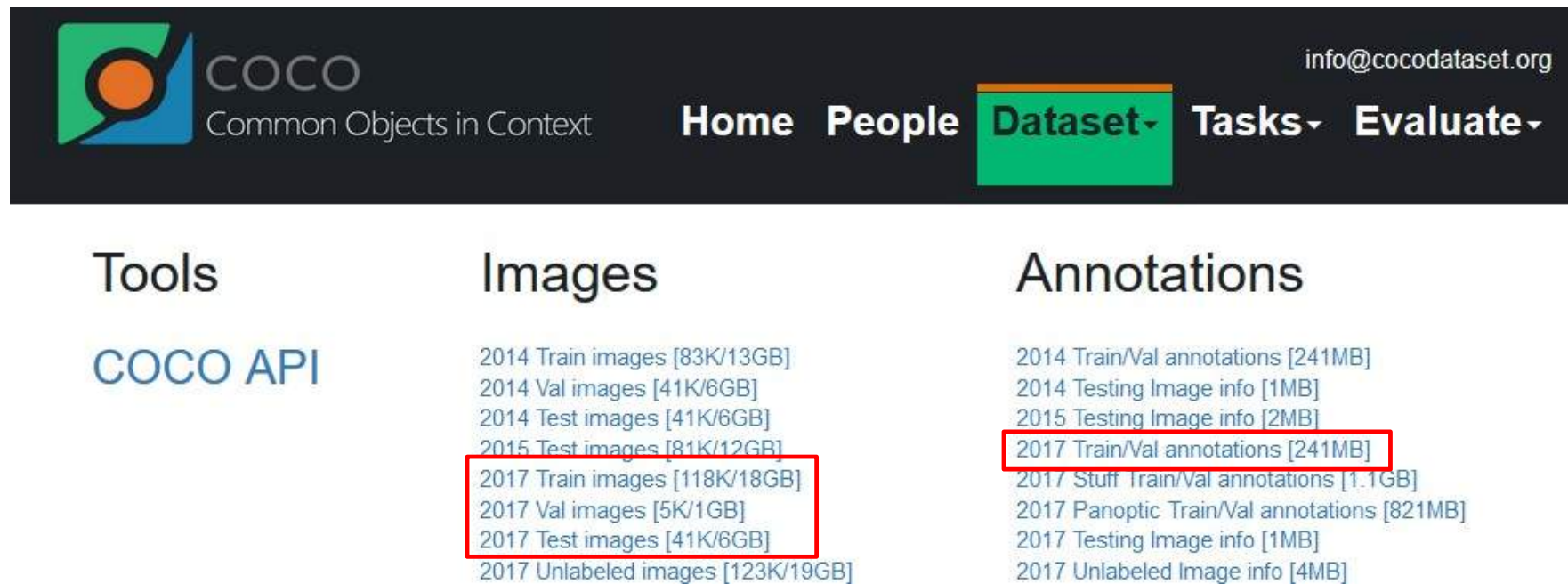
ID	COCO Paper	COCO 2017	Group
1	person	person	person
2	bicycle	bicycle	vehicle
3	car	car	vehicle
4	motorcycle	motorcycle	vehicle
5	airplane	airplane	vehicle
6	bus	bus	vehicle
7	train	train	vehicle
8	truck	truck	vehicle
9	boat	boat	vehicle
10	traffic light	traffic light	outdoor
11	fire hydrant	fire hydrant	outdoor
12	street sign	-	outdoor
13	stop sign	stop sign	outdoor
14	parking meter	parking meter	outdoor
15	bench	bench	outdoor

.....
80개의 오브젝트
카테고리

ID	COCO Paper	COCO 2017	Group
77	cell phone	cell phone	electronic
78	microwave	microwave	appliance
79	oven	oven	appliance
80	toaster	toaster	appliance
81	sink	sink	appliance
82	refrigerator	refrigerator	appliance
83	blender	-	appliance
84	book	book	indoor
85	clock	clock	indoor
86	vase	vase	indoor
87	scissors	scissors	indoor
88	teddy bear	teddy bear	indoor
89	hair drier	hair drier	indoor
90	toothbrush	toothbrush	indoor
91	hair brush	-	indoor

MS-COCO Dataset 다운로드

<http://cocodataset.org/#download>



The screenshot shows the MS-COCO Dataset website. The header includes the COCO logo (Common Objects in Context), the email info@cocodataset.org, and navigation links: Home, People, Dataset (highlighted in green), Tasks, and Evaluate. Below the header, there are three main sections: Tools, Images, and Annotations. Under Tools, the COCO API is listed. Under Images, a list of download options for 2014, 2015, and 2017 is shown. The 2017 options are highlighted with a red box: 2017 Train images [118K/18GB], 2017 Val images [5K/1GB], and 2017 Test images [41K/6GB]. Under Annotations, a list of download options for 2014, 2015, and 2017 is shown. The 2017 Train/Val annotations [241MB] is highlighted with a red box.

Tools	Images	Annotations
COCO API	2014 Train images [83K/13GB] 2014 Val images [41K/6GB] 2014 Test images [41K/6GB] 2015 Test images [81K/12GB] 2017 Train images [118K/18GB] 2017 Val images [5K/1GB] 2017 Test images [41K/6GB] 2017 Unlabeled images [123K/19GB]	2014 Train/Val annotations [241MB] 2014 Testing Image info [1MB] 2015 Testing Image info [2MB] 2017 Train/Val annotations [241MB] 2017 Stuff Train/Val annotations [1.1GB] 2017 Panoptic Train/Val annotations [821MB] 2017 Testing Image info [1MB] 2017 Unlabeled Image info [4MB]

2017년 버전이 최신 : 위에 선택된 4개 파일을 받아 사용한다

MS-COCO Explorer 기능

<http://cocodataset.org/#explorer>

The screenshot displays the MS-COCO Explorer web interface. At the top, the COCO logo (Common Objects in Context) is on the left, and the email address info@cocodataset.org is on the right. A navigation bar contains links for Home, People, Dataset (highlighted in green), Tasks, and Evaluate. A dropdown menu for the Dataset link is open, showing options: Overview, Explore (selected), Download, External, and Terms of Use. Below the navigation bar, the page title is 'COCO Explorer'. Underneath, it says 'COCO 2017 train/val browser (123,287 images, 886,284 instances). Crowd'. A large grid of icons representing various object classes is shown. The icons for 'person', 'bicycle', and 'bench' are highlighted with green boxes. At the bottom, there is a search bar with the filters 'person x', 'bicycle x', and 'bench x' applied. A 'search' button is to the right of the search bar. Below the search bar, it indicates '294 results'.

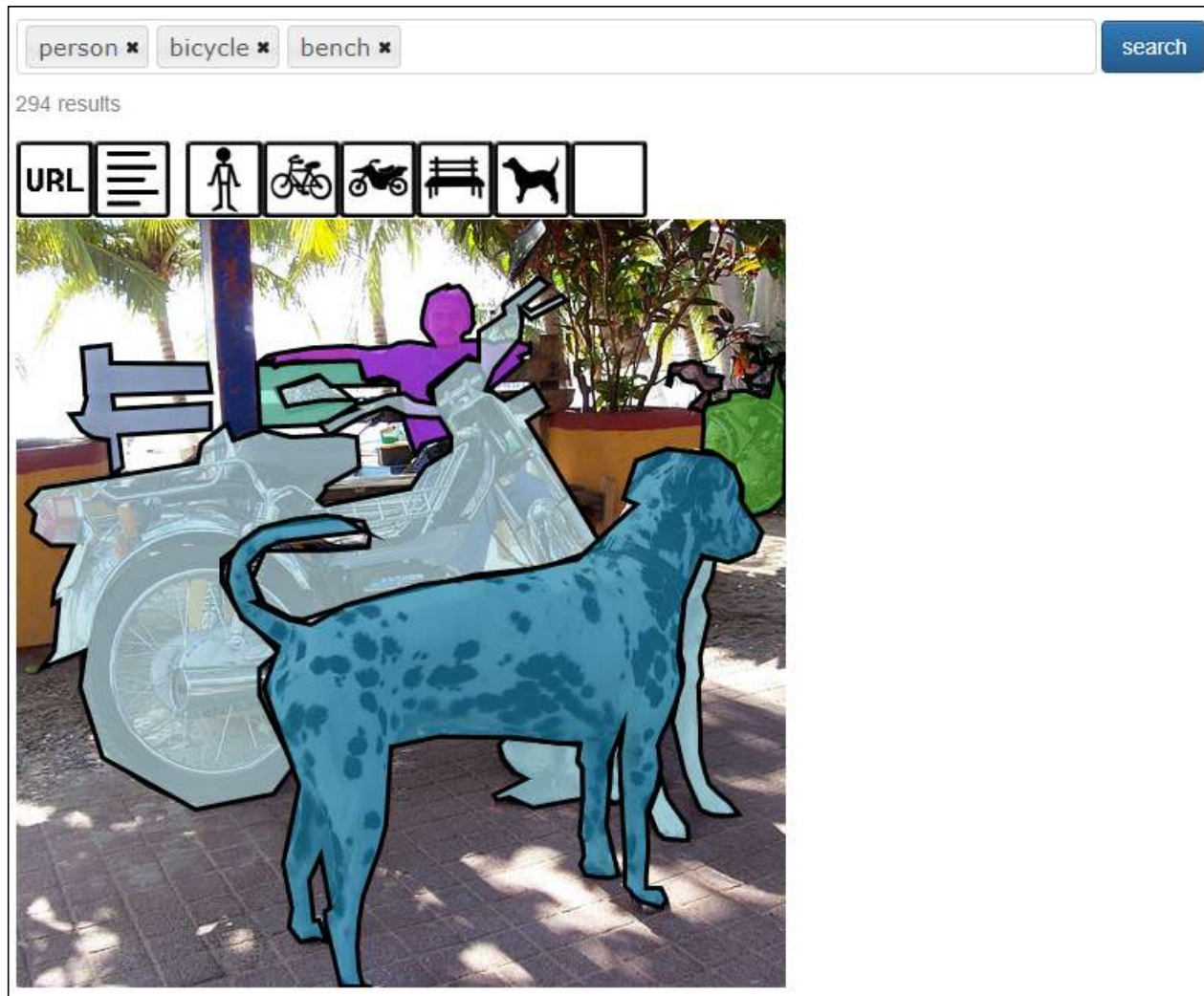
COCO Explorer

COCO 2017 train/val browser (123,287 images, 886,284 instances). Crowd

person x bicycle x bench x

294 results

MS-COCO Explorer 기능



MS-COCO Dataset 구성

COCO 2017 데이터 세트 기준

이미지 파일들

학습용 파일 **train2017**

train2017.zip

이미지명1.jpg
....
이미지명....jpg

검증용 파일 **val2017**

val2017.zip

이미지명1.jpg
....
이미지명....jpg

테스트용 파일 **test2017**

test2017.zip

이미지명1.jpg
....
이미지명....jpg

Json 파일

JSON Annotation 파일

- JSON 포맷인 한 개의 파일로 구성됨(한 라인으로 구성됨)
- 아래와 같은 대분류로 구성

info

COCO Dataset 생성 일자등을 가지는 헤더 정보임.

license

이미지 파일들의 라이선스에 대한 정보

images

모든 이미지들의 id, 파일명, 이미지 너비, 높이 정보

annotations

대상 image 및 object id
Segmentation, bounding box, 픽셀 영역등의 상세 정보

categories

80개 오브젝트 카테고리에 대한 id, 이름, Group을 가짐.

MS-COCO JSON 파일 예시

JSON 파일의 images 대분류 예시

```
"images": [  
  {  
    "license": 1,  
    "file_name": "000000037777.jpg",  
    "coco_url": "http://images.cocodataset.org/val2017/000000037777.jpg",  
    "height": 230,  
    "width": 352,  
    "date_captured": "2013-11-14 20:55:31",  
    "flickr_url": "http://farm9.staticflickr.com/8429/7839199426_f6d48aa585_z.jpg",  
    "id": 37777  
  },  
  ...  
]
```

JSON 파일의 annotations 대분류 예시

```
"annotations": [  
  {  
    "segmentation": [[510.66,423.01,511.72,420.03,...,510.45,423.01]],  
    "area": 702.1057499999998,  
    "iscrowd": 0,  
    "image_id": 289343,  
    "bbox": [473.07,395.93,38.65,28.67],  
    "category_id": 18,  
    "id": 1768  
  },  
  ...  
]
```

MS-COCO Dataset 특징

COCO Dataset은 이미지 한 개에 여러 오브젝트들을 가지고 있으며 타 데이터 세트에 비해 난이도가 높은 데이터를 제공

