

The Kubwa logo consists of the word "kubwa" in a bold, lowercase, sans-serif font, followed by a registered trademark symbol (®). The logo is centered within a white square.

kubwa®

BIGDATA & AI ANALYTICS
EXPERT COMPANY

Object Detection Data-set

| Object Detection 데이터셋

PASCAL VOC



XML Format

20개의 오브젝트 카테고리

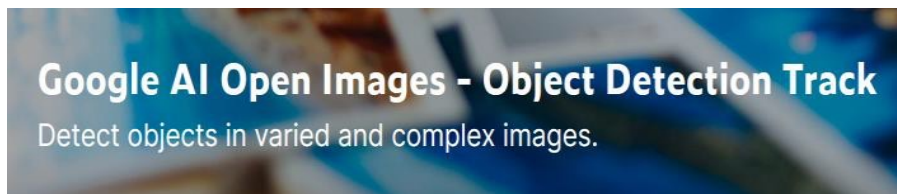
MS COCO



json Format

80개의 오브젝트 카테고리

Google Open Images



csv Format

600개의 오브젝트 카테고리

Annotation?

- Object Detection에서 이미지의 Detection 정보를 설명 파일로 제공
- Annotation은 객체(Object)의 Bounding Box 위치와 Object 이름(Class Name), 이미지 정보를 포함

VOC(.xml)

```
<annotation>
  <folder>GeneratedData_Train</folder>
  <filename>000001.png</filename>
  <path>my/path/GeneratedData_Train/000001.png</path>
  <source>
    <database>Unknown</database>
  </source>
  <size>
    <width>224</width>
    <height>224</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>21</name>
    <pose>Frontal</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <occluded>0</occluded>
    <bndbox>
      <xmin>82</xmin>
      <xmax>172</xmax>
      <ymin>88</ymin>
      <ymax>146</ymax>
    </bndbox>
  </object>
</annotation>
```

COCO(.json)

```
annotation{
  "id" : int,
  "image_id" : int,
  "category_id" : int,
  "segmentation" : RLE or [polygon],
  "area" : float,
  "bbox" : [x,y,width,height],
  "iscrowd" : 0 or 1,
}

categories[{
  "id" : int,
  "name" : str,
  "supercategory" : str,
}]
```

ImageNet(.xml)

```
<annotation>
  <filename>image-0000016.jpg</filename>
  <size>
    <width>1920</width>
    <height>1080</height>
  </size>
  <object>
    <name>sedan</name>
    <bndbox>
      <xmin>75</xmin>
      <ymin>190</ymin>
      <xmax>125</xmax>
      <ymax>210</ymax>
    </bndbox>
  </object>
</annotation>
```

Open Images(.csv)

	A	B	C	D	E	F	G	H
1	ImageID	Source	LabelName	Confidence	XMin	XMax	YMin	YMax
2	00000670070096	freiform	my077r	1	0.071905	0.145346	0.206501	0.391306
3	00000670070096	freiform	my077r	1	0.439756	0.572406	0.264153	0.435122
4	00000670070096	freiform	my077r	1	0.688455	1	0	0.552825
5	000062439959c344	freiform	my015p6	1	0.205719	0.849912	0.154144	1
6	000062439959c344	freiform	my056z4	1	0.137133	0.377634	0	0.884185
7	000064e1253d8f	freiform	my07y9	1	0	0.97305	0	0.843342
8	000064e1253d8f	freiform	my0k4j	1	0	0.513534	0.321356	0.886601
9	000064e1253d8f	freiform	my0k4j	1	0.016515	0.260220	0.295368	0.462906
10	000064e1253d8f	freiform	my0k4j	1	0.481448	0.904378	0.232029	0.489117
11	000064e1253d8f	freiform	my0k4j	1	0.751517	1	0.33235	0.658053
12	000132c20b84209b	freiform	my01g311	1	0.321824	0.5701	0.157304	1
13	000132c20b84209b	freiform	my0283cl	1	0.154215	0.759208	0.258625	0.394532
14	000132c20b84209b	freiform	my03q69	1	0.13551	0.231280	0.266886	0.365227
15	000132c20b84209b	freiform	my03q69	1	0.430795	0.510488	0.160632	0.235236
16	000132c20b84209b	freiform	my03q69	1	0.565238	0.69831	0.171936	0.349402
17	000132c20b84209b	freiform	my03q69	1	0.694072	0.850020	0.258973	0.442091
18	000132c20b84209b	freiform	my04hg8	1	0.135425	0.229449	0.270051	0.416875
19	000132c20b84209b	freiform	my04hg8	1	0.440367	0.516296	0.163887	0.301875
20	000132c20b84209b	freiform	my04hg8	1	0.581826	0.654673	0.174051	0.319746
21	000132c20b84209b	freiform	my04hg8	1	0.694485	0.780885	0.263275	0.420203
22	000132c20b84209b	freiform	my05s655	1	0.033341	0.28196	0.270638	1
23	000132c20b84209b	freiform	my05s655	1	0.567332	0.844056	0.167807	1
24	000132c20b84209b	freiform	my05s655	1	0.658132	0.85297	0.358146	1
25	000132c20b84209b	freiform	my0992d	1	0.040982	0.283854	0.384433	1
26	000132c20b84209b	freiform	my0992d	1	0.043871	0.878808	0.266582	1
27	000132c20b84209b	freiform	my0992d	1	0.332815	0.571821	0.272744	0.945032

| Annotation?

- Object Detection에서 이미지의 Detection 정보를 설명 파일로 제공
- Annotation은 객체(Object)의 Bounding Box 위치와 Object 이름(Class Name), 이미지 정보를 포함
 - Class index = Category index = Class name
 - Bounding Box 좌표
(x_min, x_max, y_min, y_max) or (x_center, y_center, width, height)

“annotation을 제작한다” 또는 “labeling 한다”

| Annotation?

- Object Detection에서 이미지의 Detection 정보를 설명 파일로 제공
- Annotation은 객체(Object)의 Bounding Box 위치와 Object 이름(Class Name), 이미지 정보를 포함

이미지/영상 데이터셋 종류

- PASCAL VOC: <http://host.robots.ox.ac.uk/pascal/VOC/>
- MS COCO: <https://cocodataset.org/#home>
- Google OpenImages: <https://github.com/openimages/dataset>
- Udacity: <https://github.com/udacity/self-driving-car/tree/master/annotations>
- KITTI 2D: <http://www.cvlibs.net/datasets/kitti/>



BIGDATA & AI ANALYTICS
EXPERT COMPANY

1. PASCAL VOC

PASCAL VOC 2012

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



Visual Object Classes Challenge 2012 (VOC2012)

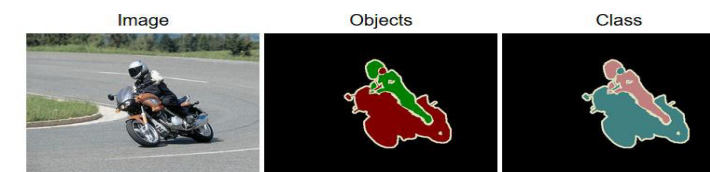


[click on an image to see the annotation]

Classification /Detection



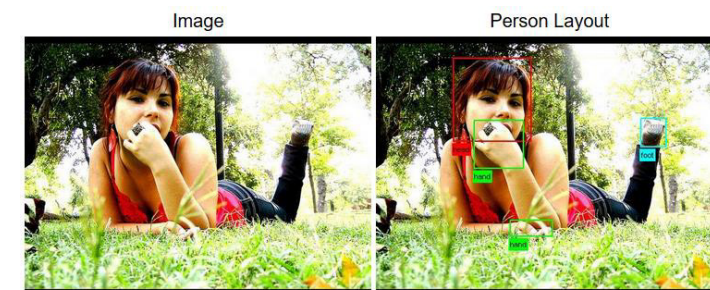
Segmentation



Action Classification



Person Layout



VOC20XX

- ├ Annotations
- ├ ImageSets
- ├ JPEGImages
- ├ SegmentationClass
- └ SegmentationObject

- Annotations: JPEGImages 폴더 속 원본 이미지와 같은 이름들의 xml파일로 구성
- ImageSets: 어떤 이미지 그룹을 test, train, trainval, val로 사용할 것인지, 특정 클래스 정보를 포함
- JPEGImages: *.jpg확장자를 가진 이미지 파일들이 모여있는 폴더로 Object Detection에서 입력 데이터
- SegmentationClass: Semantic segmentation을 학습하기 위한 label 이미지
- SegmentationObject: Instance segmentation을 학습하기 위한 label 이미지

*. xml 형식을 parsing 하여 사용

〈size〉 : xml파일과 대응되는 이미지의 width, height, channels 정보에 대한 tag

〈width〉 : xml파일에 대응되는 이미지의 width 값

〈height〉 : xml파일에 대응되는 이미지의 height 값

〈depth〉 : xml파일에 대응되는 이미지의 channels 값

〈object〉 : xml파일과 대응되는 이미지속에 object의 정보에 대한 tag

〈name〉 : 클래스 이름을 의미

〈bndbox〉 : 해당 object의 바운딩상자의 정보에 대한 tag

xmin : object 바운딩상자의 왼쪽상단의 x축 좌표값

ymin : object 바운딩상자의 왼쪽상단의 y축 좌표값

xmax : object 바운딩상자의 우측하단의 x축 좌표값

ymax : object 바운딩상자의 우측하단의 y축 좌표값

Bounding Box 좌표

- xmin : object 바운딩 상자의 왼쪽상단의 x축 좌표값
- ymin : object 바운딩 상자의 왼쪽상단의 y축 좌표값
- xmax : object 바운딩 상자의 우측하단의 x축 좌표값
- ymax : object 바운딩 상자의 우측하단의 y축 좌표값



XML 파일 로드 파싱 순서

1. 입력된 xml 파일 위치를 받아 xml 파일을 open
2. open한 xml파일을 이용하여 xml.etree.ElementTree 를 사용하여 파싱
3. xml.etree.ElementTree 객체로 부터 xml의 루트(여기에서는 tag)를 가져오기
4. <size> tag를 찾는다
5. 찾은 <size> tag에서 width, height, depth tag를 찾아서 값들을 불러오기
6. xml의 루트에서 모든 <objects> tag를 찾는다
7. 여러 <objects> tags를
순회하면서 <name>, <bndbox>, <xmin>, <ymin>, <xmax>, <ymax> tag를 찾아 값들을 불러온다.

Annotation 파일에서 얻은 Label 정보를 이용하여 이미지에 Object Detection box 표현하기

Object Detection label 정보 표현 절차

1. Annotations폴더의 파일 리스트를 불러옴
2. JPEGImages폴더에서 XML 파일과 대응되는 이미지를 찾음
3. 파일 리스트를 순회를 돌면서 파일에서 label 정보를 load
4. load된 label 정보를 이용하여 Object Detection 정보를 시각화

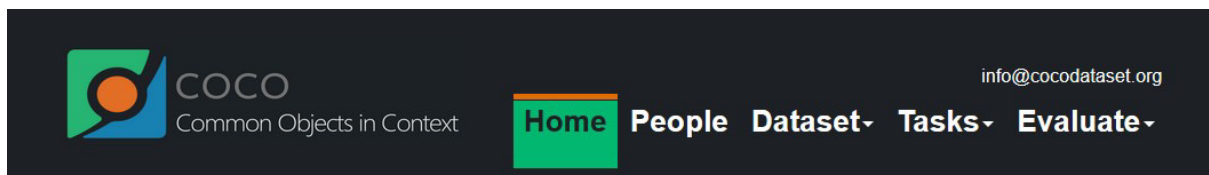
The Kubwa logo is centered within a white square. It consists of the word "kubwa" in a bold, lowercase, sans-serif font, followed by a registered trademark symbol (®).

kubwa®

BIGDATA & AI ANALYTICS
EXPERT COMPANY

2. MS COCO

MS COCO Dataset



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

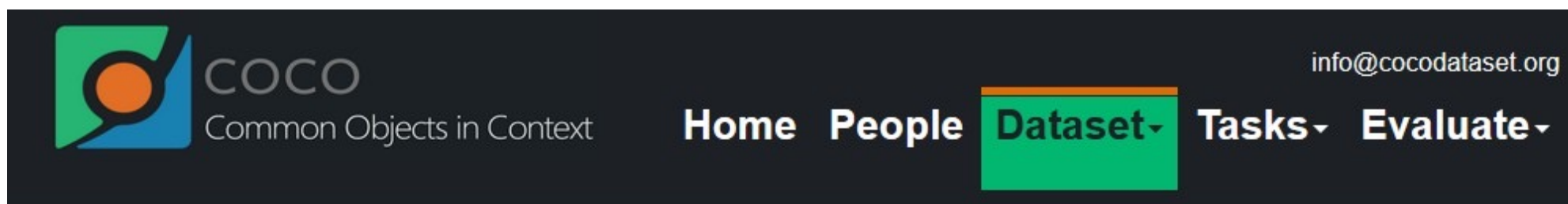
Sponsors



- 80개 Object Category
- 300K의 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 Pretrained된 모델을 제공함

COCO Dataset 다운로드

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



Tools

COCO API

Images

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]

Annotations

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]

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 정보

license

이미지 파일들의 라이선스에 대한 정보

images

모든 이미지들의 id, 파일명, 이미지 너비, 높이 정보

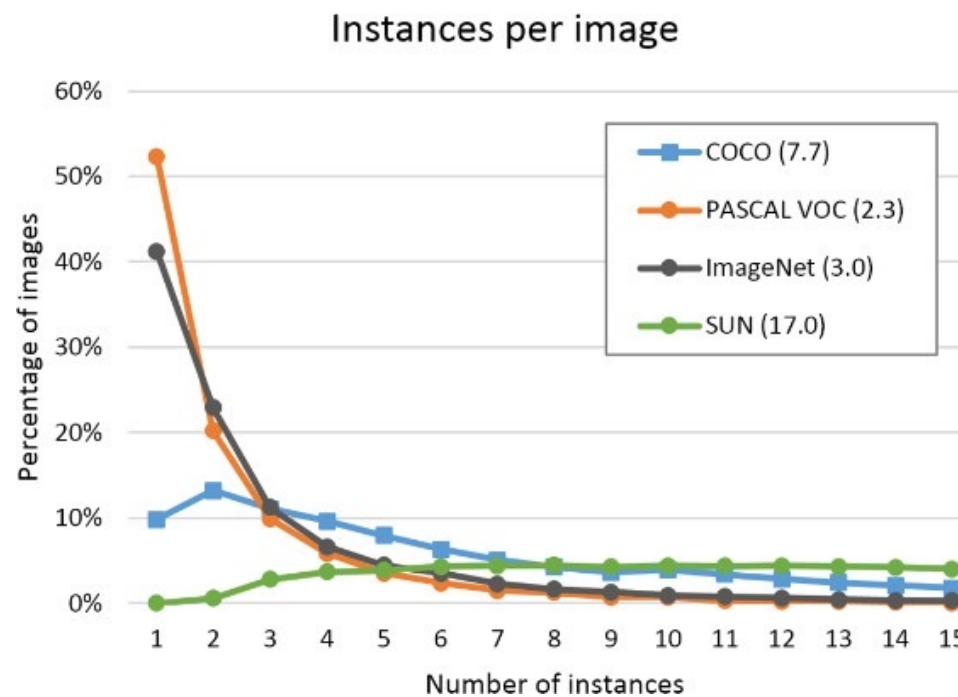
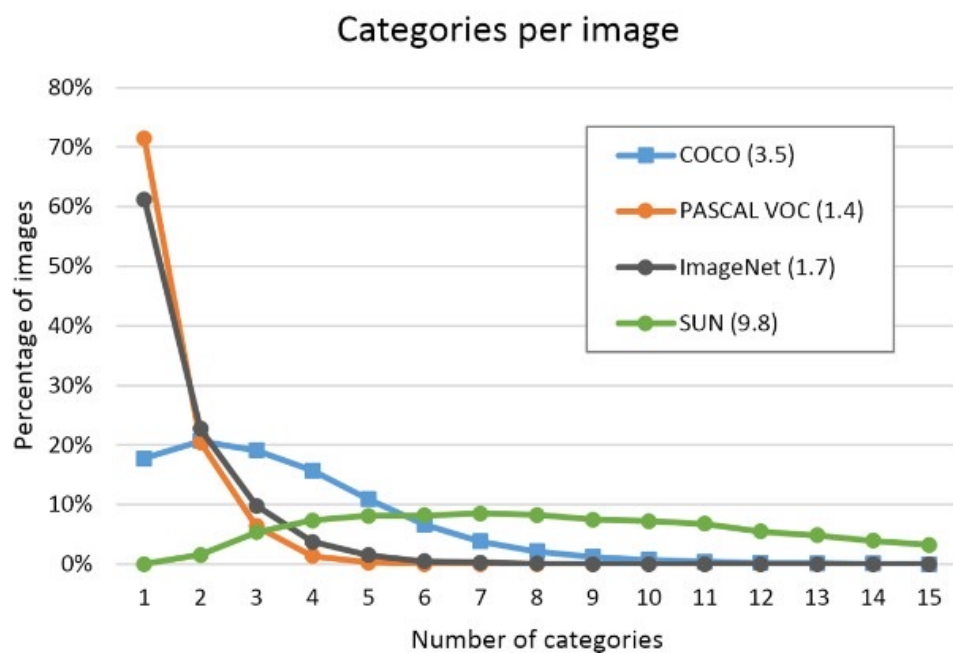
annotations

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

categories

80개 오브젝트 카테고리에 대한 id, 이름, Group 정보

- COCO Dataset은 하나의 이미지 데이터셋에 복수의 오브젝트들로 구성되어 타 데이터 세트에 비해 난이도가 높은 데이터를 제공



JSON 파일의 Categories 예시

```
Categories
[
  {
    "id": int,
    "name": str,
    "supercategory": str,
  }
]
```

```
"categories":
[
  { "id": 1, "name": "rose", "supercategory": "flower" },
  { "id": 2, "name": "tulip", "supercategory": "flower" },
  { "id": 10, "name": "Apple", "supercategory": "fruit" }
]
```

JSON 파일의 Images

```
image{
  "id": int,
  "width": int,
  "height": int,
  "file_name": str,
  "license": int,
  "flickr_url": str,
  "coco_url": str,
  "date_captured": datetime,
}
```

```
"images": [
  {
    "id": 397133,
    "width": 640,
    "height": 640,
    "file_name": "101.jpg",
    "license": 1,
    "date_captured": "2019-12-04 17:02:52"
  },
  {
    "id": 397122,
    "height": 640,
    "width": 640,
    "file_name": "102.jpg",
    "license": 1,
    "date_captured": "2019-12-04 17:02:52"
  }
]
```

JSON 파일의 Annotations 예시

```
annotation{  
  "id": int,  
  "image_id": int,  
  "category_id": int,  
  "segmentation": RLE or [polygon],  
  "area": float,  
  "bbox": [x,y,width,height],  
  "iscrowd": 0 or 1,  
}  
  
"annotations": [  
  {  
    "segmentation":  
      [[510.66,423.01,511.72,420.03,...,510.45,423.01]],  
    "area": 702.10,  
    "iscrowd": 0,  
    "image_id": 397133,  
    "bbox": [433.07,355.93,138.65,228.67],  
    "category_id": 18,  
    "id": 1768  
  },  
  {  
    "segmentation":  
    {  
      "counts": [12,56,198,10]  
      "size": [120, 240]  
    }  
    "area": 500.2,  
    "iscrowd": 1,  
    "image_id": 397122,  
    "bbox": [473.07,395.93,38.65,28.67],  
    "category_id": 18,  
    "id": 1768  
  }  
]
```



80개의 Object Category

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

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

Coco 클래스 id와 name 매핑

80개의 Object Category

OpenCV DNN 	Tensorflow Faster R-CNN 모델 로드	0 ~ 90 까지
	Tensorflow SSD 모델 로드	1 ~ 91 까지
	Tensorflow Mask R-CNN 모델 로드	0 ~ 90 까지
	Darknet YOLO 모델 로드	0 ~ 79 까지
Tensorflow  TensorFlow		1 ~ 91까지

The Kubwa logo consists of the word "kubwa" in a bold, lowercase, sans-serif font, followed by a registered trademark symbol (®). The logo is centered within a white square.

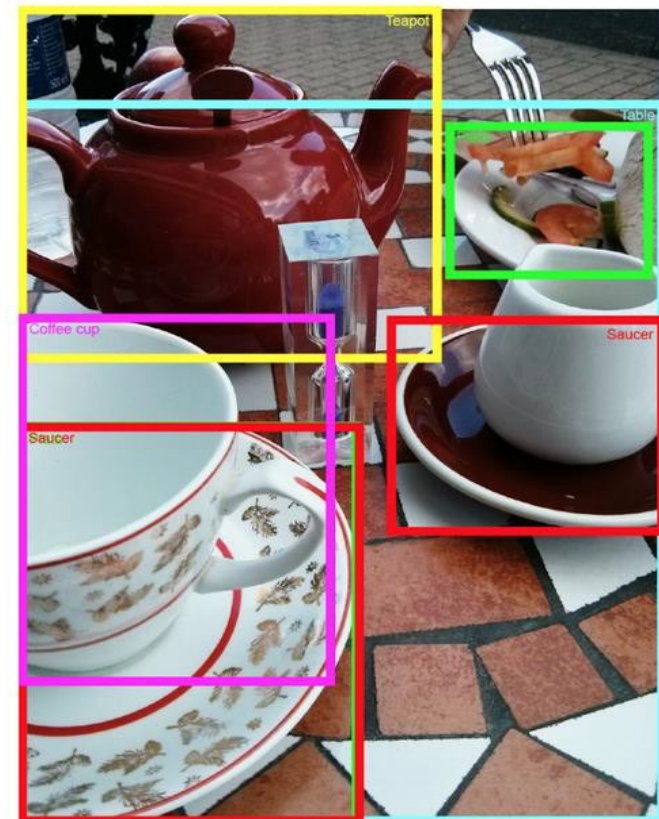
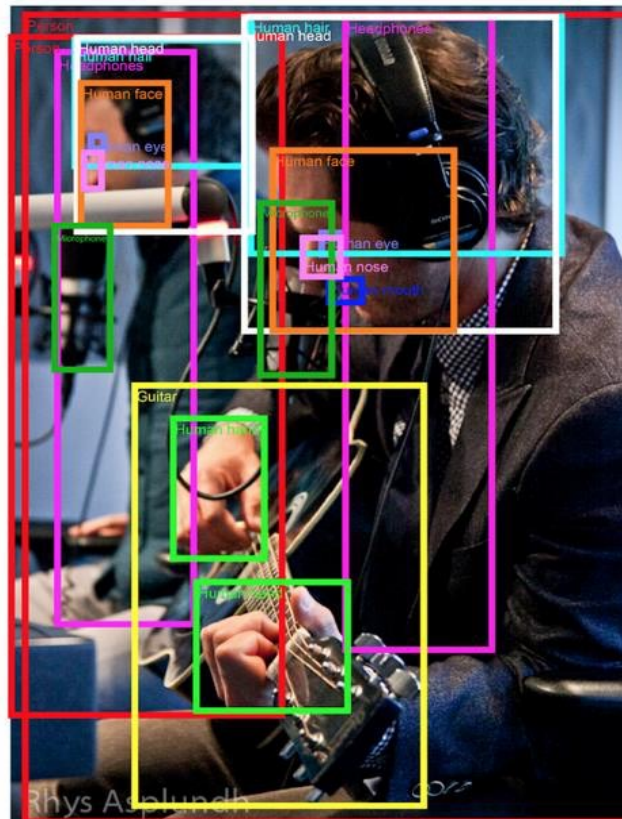
kubwa®

BIGDATA & AI ANALYTICS
EXPERT COMPANY

3. Google Open Image

600개의 오브젝트 카테고리

- 1,743,042 training images
- 41,620 validation images
- 125,436 test images



| Open Image Dataset V6 +

<https://storage.googleapis.com/openimages/web/index.html>

The screenshot shows the homepage of the Open Images Dataset V6 + Extensions. The background is a collage of images with various bounding boxes and labels. The main title 'Open Images Dataset V6 + Extensions' is prominently displayed in white text. Below the title, several statistics are listed in a light gray font. At the bottom, there are seven white buttons with black text arranged in two rows.

Open Images Dataset V6 + Extensions

15,851,536 boxes on 600 categories
2,785,498 instance segmentations on 350 categories
3,284,282 relationship annotations on 1,466 relationships
507,444 localized narratives
59,919,574 image-level labels on 19,957 categories
Extension - 478,000 crowdsourced images with 6,000+ categories

Explore Description Download
Extended News Extras Challenge

Google Open Image 특징

- 600개의 Object Class 지원
- 350개의 Instance Segmentation 지원

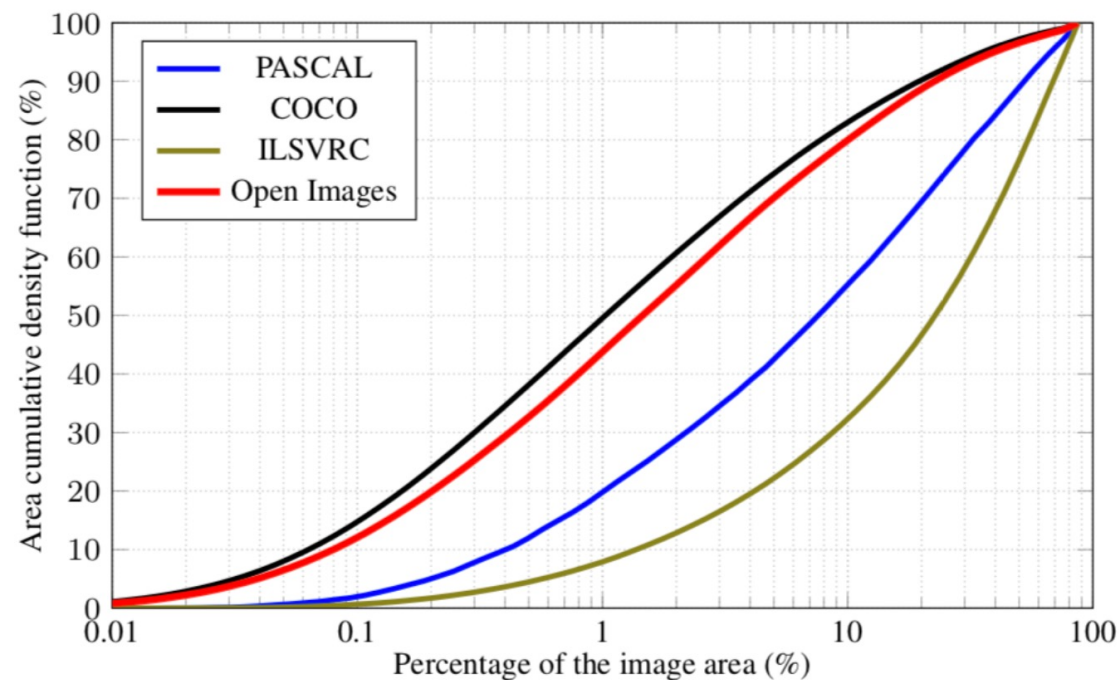
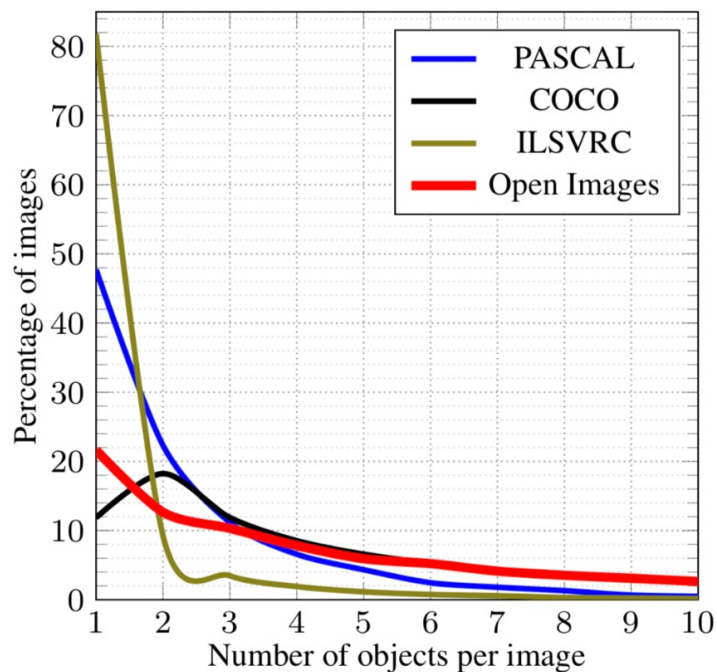


Image-level labels

	Train	Validation	Test	# Classes	# Trainable Classes
Images	9,011,219	41,620	125,436	-	-
Machine-Generated Labels	164,819,642	681,179	2,061,177	15,387	9,034
Human-Verified Labels	57,524,352 pos: 19,856,086 neg: 37,668,266	595,339 pos: 367,263 neg: 228,076	1,799,883 pos: 1,110,124 neg: 689,759	19,957	9,605

Bounding boxes

	Train	Validation	Test	# Classes
Images	1,743,042	41,620	125,436	-
Boxes	14,610,229	303,980	937,327	600

Object segmentations

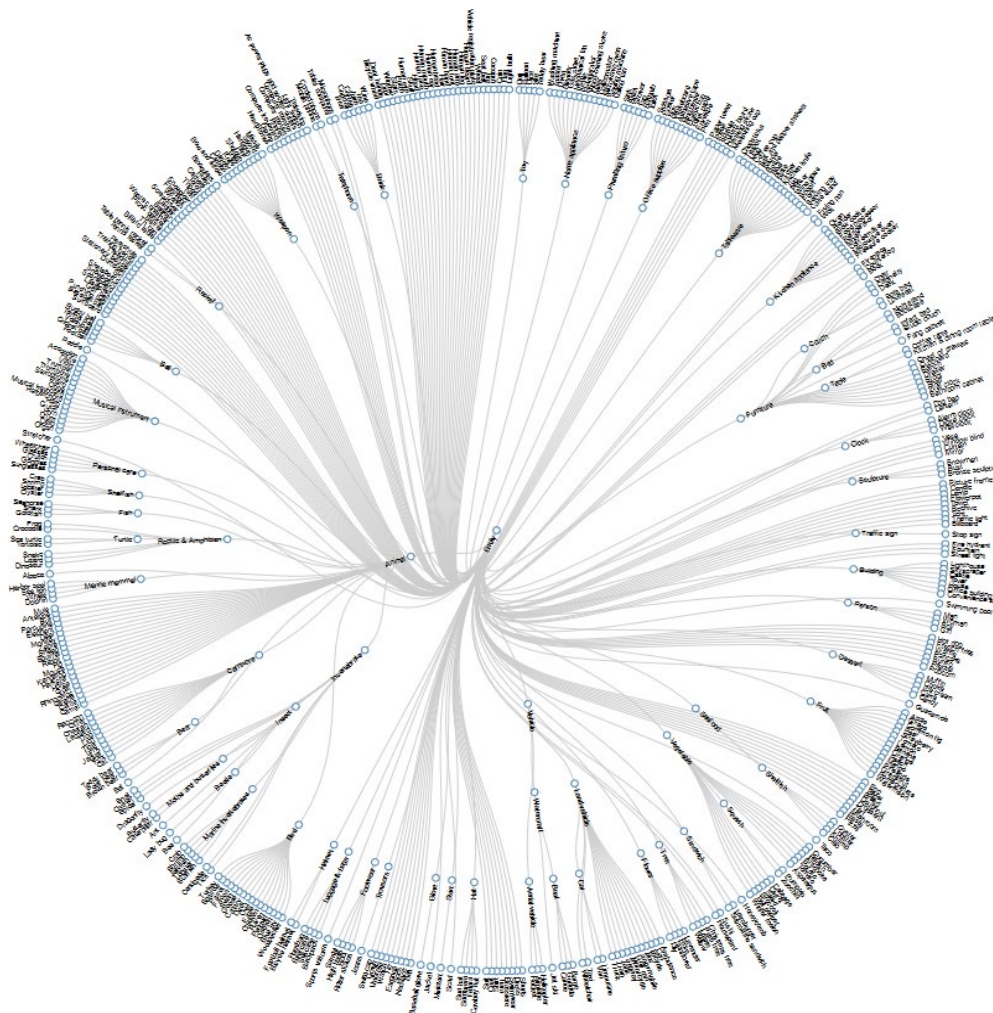
	Train	Validation	Test	# Classes
Images	944,037	13,524	40,386	-
Instance masks	2,686,666	24,730	74,102	350

Visual relationships

	Train	Validation	Test	# Distinct relationship triplets	# Classes
Relationship triplets	3,174,293 non-attr: 348,560	27,243 non-attr: 4,951	82,746 non-attr: 14,403	1,466 non-attr: 1,384	288

Google Open Image 구조

https://storage.googleapis.com/openimages/challenge_2018/bbox_labels_500_hierarchy_visualizer/circle.html



The Kubwa logo is centered within a white square. It consists of the word "kubwa" in a bold, lowercase, sans-serif font, followed by a registered trademark symbol (®).

kubwa®

BIGDATA & AI ANALYTICS
EXPERT COMPANY

4. Annotation 실습