

ImageNet

▼ Format	PASCAL VOC
# class	1000
# test	100000
# train	1281167
# val	50000
≡ 사이트	https://www.image-net.org/index.php
≡ 태그	

About

ImageNet은 **ImageNet Large Scale Visual Recognition Challenge (ILSVRC)**를 위해 만들어진 Dataset

WordNet이라는 heirachy를 따르는데, WordNet은 여러개의 단어, 구문으로 이루어진 **synset**을 의미

100,000개의 synsets로 구성됨(그 중 80,000개 이상이 명사)

각 synsets당 1000개 이상의 이미지를 제공

1,000개의 class (class ≠ synset)

```
n01440764 tench, Tinca tinca
n01443537 goldfish, Carassius auratus
n01484850 great white shark, white shark, man-eater, man-eating shark, Carcharodon carcharias
n01491361 tiger shark, Galeocerdo cuvieri
n01494475 hammerhead, hammerhead shark
n01496331 electric ray, crampfish, numbfish, torpedo
n01498041 stingray
n01514668 cock
n01514859 hen
n01518878 ostrich, Struthio camelus
n01530575 brambling, Fringilla montifringilla
n01531178 goldfinch, Carduelis carduelis
n01532829 house finch, linnet, Carpodacus mexicanus
n01534433 junco, snowbird
n01537544 indigo bunting, indigo finch, indigo bird, Passerina cyanea
n01558993 robin, American robin, Turdus migratorius
n01560419 bulbul
n01580077 jay
n01582220 magpie
n01592084 chickadee
n01601694 water ouzel, dipper
n01608432 kite
n01614925 bald eagle, American eagle, Haliaeetus leucocephalus
n01616318 vulture
n01622779 great grey owl, great gray owl, Strix nebulosa
n01629819 European fire salamander, Salamandra salamandra
n01630670 common newt, Triturus vulgaris
n01631663 eft
n01632458 spotted salamander, Ambystoma maculatum
n01632777 axolotl, mud puppy, Ambystoma mexicanum
n01641577 bullfrog, Rana catesbeiana
n01644373 tree frog, tree-frog
n01644900 tailed frog, bell toad, ribbed toad, tailed toad, Ascaphus trui
n01664065 loggerhead, loggerhead turtle, Caretta caretta
```

LOC_synset_mapping.txt


1,281,167개의 train data

50,000개의 validation data

100,000개의 test data

구조

Object Detection dataset 중 하나인 **PASCAL VOC dataset**의 구조

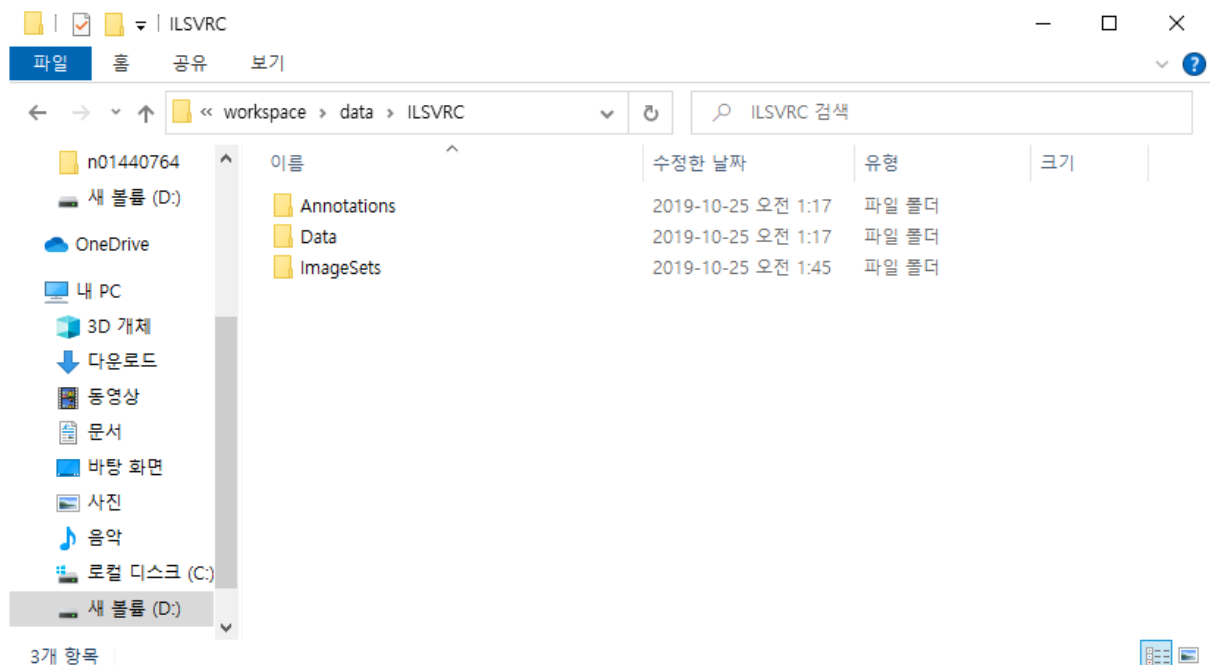
 참고 사이트

[PASCAL VOC dataset의 구조](#)

폴더 구조

```
VOC20XX
├─ Annotations
├─ ImageSets
├─ JPEGImages
├─ SegmentationClass
└─ SegmentationObject
```

- **Annotations** : **JPEGImages** 폴더 속 원본 이미지와 같은 이름들의 **xml** 파일들이 존재합니다. Object Detection을 위한 **정답 데이터** 이 됩니다.
- **ImageSets** : 어떤 이미지 그룹을 **test** , **train** , **trainval** , **val** 로 사용할 것인지, 특정 클래스가 어떤 이미지에 있는지 등에 대한 정보들을 포함하고 있는 폴더입니다.
- **JPEGImages** : **.jpg** 확장자를 가진 이미지 파일들이 모여있는 폴더입니다. Object Detection에서 **입력 데이터** 가 됩니다.
- **SegmentationClass** : Semantic segmentation을 학습하기 위한 **label** 이미지입니다.
- **SegmentationObject** : Instance segmentation을 학습하기 위한 **label** 이미지입니다.



ImageNet의 압축을 풀면

→ ILSVRC가 Segmentation task는 다루지 않는 competition이라 Segmentation 관련 폴더는 없음

XML 구조

- `<size>` : xml파일과 대응되는 이미지의 width, height, channels 정보에 대한 tag입니다.
 - `<width>` : xml파일에 대응되는 이미지의 width값입니다.
 - `<height>` : xml파일에 대응되는 이미지의 height값입니다.
 - `<depth>` : xml파일에 대응되는 이미지의 channels값입니다.
- `<object>` : xml파일과 대응되는 이미지속에 object의 정보에 대한 tag입니다.
 - `<name>` : 클래스 이름을 의미합니다.
 - `<bndbox>` : 해당 object의 바운딩상자의 정보에 대한 tag입니다.
 - `xmin` : object 바운딩상자의 왼쪽상단의 x축 좌표값입니다.
 - `ymin` : object 바운딩상자의 왼쪽상단의 y축 좌표값입니다.
 - `xmax` : object 바운딩상자의 우측하단의 x축 좌표값입니다.
 - `ymax` : object 바운딩상자의 우측하단의 y축 좌표값입니다.

```
<annotation>
  <folder>n01443537</folder>
  <filename>n01443537_129</filename>
  <source>
    <database>ILSVRC_2012</database>
  </source>
  <size>
    <width>490</width>
    <height>322</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>n01443537</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>107</xmin>
      <ymin>15</ymin>
      <xmax>363</xmax>
      <ymax>147</ymax>
    </bndbox>
  </object>
</object>
```

```
<name>n01443537</name>
<pose>Unspecified</pose>
<truncated>0</truncated>
<difficult>0</difficult>
<bndbox>
  <xmin>212</xmin>
  <ymin>51</ymin>
  <xmax>379</xmax>
  <ymax>216</ymax>
</bndbox>
</object>
</annotation>
```

→ Object Detection도 다루다 보니 한 이미지에 여러 object가 들어가 있을 수 있음 (EDA 필요)

🍎한 이미지에 서로 다른 class의 object가 동시에 들어 있는 이미지가 존재

→ boundary box로 이미지를 crop해서 training image로 사용해야 함

🍎한 이미지에는 같은 class의 object만 존재

→ 선택해야함 crop해서 할지 전체 이미지를 사용할지

🍏test 이미지에 대한 annotation은 제공을 안함

→ 학생/organization 메일로 download 신청을 해야만 받을 수 있음

→ 현재 받은건 kaggle에서 배포하는 거라 test 이미지에 대한 정답 값이 없음

→ 회사 organization 메일로 신청하거나 train dataset을 직접 나눠서 test 용으로 만들어야 함

🍏test 이미지에 한개의 object가 있다는 보장?