



Πολυτεχνείο
Κρήτης

Σχολή Ηλεκτρολόγων
Μηχανικών & Μηχανικών
Υπολογιστών

Data description

Name: [Stanford Drone Dataset](#)

It consists of eight unique scenes, with the following classes being available to detect:

- 1) Pedestrian (Renamed Person)
- 2) Biker
- 3) Skater (Merged and renamed Person)
- 4) Car
- 5) Bus
- 6) Cart (Merged and renamed Car)

Therefore, for our project we have four classes.

Person (22673 annotations)

Biker (11479 annotations)

Bus (101 annotations)

Car (1512 annotations)

Anchor Boxes Parameters: sizes = {16, 32, 64, 128, 256}, aspect ratios = {1:2, 1:1, 2:1, 3:1}

(16x32, 16x16, 32x16, 64x16

32x64, 32x32, 64x32, 128x32

64x128, 64x64, 128x64, 256x64

128x256, 128x128, 256x128

256x256)

Total 16 possible anchor boxes.

Dataset Format: Pascal VOC (path/to/image, x1,y1,x2,y2,class_name)