

---

The list is now **archived**. Please see these fantastic resources for more recent datasets: [satellite-image-deep-learning](#) & [Awesome\\_Satellite\\_Benchmark\\_Datasets](#)

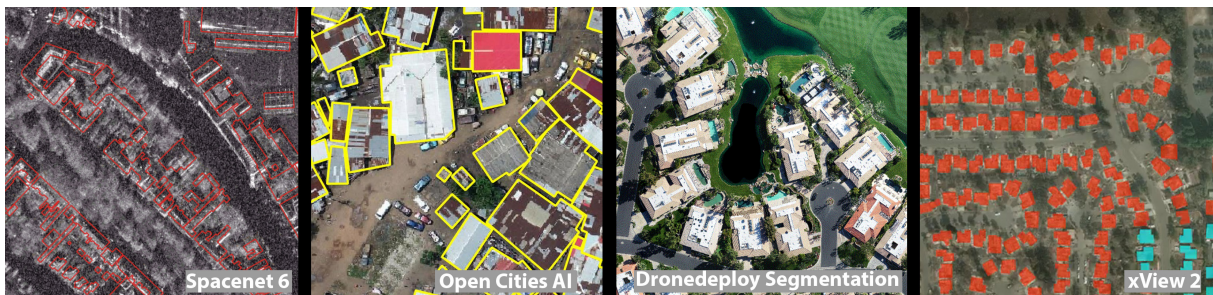
## Awesome Satellite Imagery Datasets

List of aerial and satellite imagery datasets with annotations for computer vision and deep learning. Newest datasets at the top of each category (Instance segmentation, object detection, semantic segmentation, scene classification, other).



### Recent additions and ongoing competitions

- **FloodNet** (*University of Maryland, Jun 2021*)  
2343 image chips (drone imagery), 10 landcover categories (background, water, building flooded, building non-flooded, road-flooded, ...). Paper: Rahnemoonfar et al., 2021
- **PASTIS : Panoptic Agricultural Satellite Time Series** (*IGN, July 2021*)  
124,422 Agricultural parcels, 2,433 Sentinel-2 image chip timeseries, France, panoptic labels (instance index + semantic label for each pixel). Paper: Garnot & Landrieu 2021
- **xView3 Dark Vessel Detection 2021** (*xView3 Team, Aug 2021*)  
Maritime object bounding boxes for 1k Sentinel-1 scenes (VH & VV polarizations), ancillary data (land/ice mask, bathymetry, wind speed, direction, quality).



---

## 1. Instance Segmentation

- **PASTIS: Panoptic Agricultural Satellite Time Series** (*IGN, July 2021*)  
124,422 Agricultural parcels, 2,433 Sentinel-2 image chip timeseries, France, panoptic labels (instance index + semantic label for each pixel). Paper: Garnot & Landrieu 2021
- **SpaceNet 7: Multi-Temporal Urban Development Challenge** (*CosmiQ Works, Planet, Aug 2020*)  
Monthly building footprints and Planet imagery (4m. res) timeseries for 2 years, 100 locations around the globe, for building footprint evolution & address propagation.
- **RarePlanes: Synthetic Data Takes Flight** (*CosmiQ Works, A.I.Reverie, June 2020*)  
Synthetic (630k planes, 50k images) and real (14.7k planes, 253 Worldview-3 images (0.3m res.), 122 locations, 22 countries) plane annotations & properties and satellite images. Tools. Paper: Shermeyer et al. 2020
- **SpaceNet: Multi-Sensor All-Weather Mapping** (*CosmiQ Works, Capella Space, Maxar, AWS, Intel, Feb 2020*)  
48k building footprints (enhanced 3DBAG dataset, building height attributes), Capella Space SAR data (0.5m res., four polarizations) & Worldview-3 imagery (0.3m res.), Rotterdam, Netherlands.
- **Agriculture-Vision Database & CVPR 2020 challenge** (*UIUC, Intelinair, CVPR, Jan 2020*)  
Agricultural Pattern Analysis, 21k aerial farmland images (RGB-NIR, USA, 2019 season, 512x512px chips), label masks for 6 field anomaly patterns (Cloud shadow, Double plant, Planter skip, Standing Water, Waterway and Weed cluster). Paper: Chiu et al. 2020
- **iSAID: Large-scale Dataset for Object Detection in Aerial Images** (*IIT & Wuhan University, Dec 2019*)  
15 categories from plane to bridge, 188k instances, object instances and segmentation masks (MS COCO format), Google Earth & JL-1 image chips, Faster-RCNN baseline model (MXNet), devkit, Academic use only, replaces DOTA dataset, Paper: Zamir et al. 2019
- **xView 2 Building Damage Assessment Challenge** (*DIUx, Nov 2019*) .  
550k building footprints & 4 damage scale categories, 20 global locations and 7 disaster types (wildfire, landslides, dam collapses, volcanic eruptions, earthquakes/tsunamis, wind, flooding), Worldview-3 imagery (0.3m res.), pre-trained baseline model. Paper: Gupta et al. 2019
- **Microsoft BuildingFootprints Canada & USA & Uganda/Tanzania & Australia** (*Microsoft, Mar 2019*)  
12.6mil (Canada) & 125.2mil (USA) & 17.9mil (Uganda/Tanzania) & 11.3mil (Australia) building footprints, GeoJSON format, delineation based on Bing imagery using ResNet34 architecture.

- 
- **SpaceNet 4: Off-Nadir Buildings** (*CosmiQ Works, DigitalGlobe, Radiant Solutions, AWS, Dec 2018*)  
126k building footprints (Atlanta), 27 WorldView 2 images (0.3m res.) from 7-54 degrees off-nadir angle. Bi-cubically resampled to same number of pixels in each image to counter coarser native resolution with higher off-nadir angles, Paper: Weir et al. 2019
  - **Airbus Ship Detection Challenge** (*Airbus, Nov 2018*)  
131k ships, 104k train / 88k test image chips, satellite imagery (1.5m res.), raster mask labels in run-length encoding format, Kaggle kernels.
  - **Open AI Challenge: Tanzania** (*WeRobotics & Wordlbank, Nov 2018*)  
Building footprints & 3 building conditions, RGB UAV imagery - [Link to data](#)
  - **LPIS agricultural field boundaries** Denmark - Netherlands - France  
Annual datasets. Denmark: 293 crop/vegetation categories, 600k parcels. Netherlands: 294 crop/vegetation categories, 780k parcels
  - **CrowdAI Mapping Challenge** (*Humanity & Inclusion NGO, May 2018*)  
Buildings footprints, RGB satellite imagery, COCO data format
  - **SpaceNet 2: Building Detection v2** (*CosmiQ Works, Radiant Solutions, NVIDIA, May 2017*)  
685k building footprints, 3/8band Worldview-3 imagery (0.3m res.), 5 cities, SpaceNet Challenge Asset Library
  - **SpaceNet 1: Building Detection v1** (*CosmiQ Works, Radiant Solutions, NVIDIA, Jan 2017*)  
Building footprints (Rio de Janeiro), 3/8band Worldview-3 imagery (0.5m res.), SpaceNet Challenge Asset Library

## 2. Object Detection

- **Airbus Aircraft Detection** (*Airbus, Mar 2021*)  
Aircraft bounding boxes, 103 images of worldwide airports (Pleiades, 0.5m res., 2560px).
- **Airbus Oil Storage Detection** (*Airbus, Mar 2021*)  
Oil storage tank annotations, 98 worldwide images (SPOT, 1.2m res., 2560px).
- **xView3 Dark Vessel Detection 2021** (*xView3 Team, Aug 2021*)  
Maritime object bounding boxes for 1k Sentinel-1 scenes (VH & VV polarizations), ancillary data (land/ice mask, bathymetry, wind speed, direction, quality).
- **AFO - Aerial dataset of floating objects** (*Gašienica-Józkowy et al, Jun 2020*)  
3647 drone images from 50 scenes, 39991 objects with 6 categories (human, wind/sup-board, boat, buoy, sailboat, kayak), Darknet YOLO format, Paper: Authors: Gašienica-Józkowy et al. 2021

- 
- **NEON Tree Crowns Dataset** (*Weinstein et al., 2020*)  
Individual tree crown objects, height&area estimates, 100 million instances, 37 geographic sites across the US, DeepForest Python package, Paper: Weinstein et al. 2020
  - **xView 2018 Detection Challenge** (*DIUx, Jul 2018*)  
60 categories from helicopter to stadium, 1 million instances, Worldview-3 imagery (0.3m res.), COCO data format, pre-trained Tensorflow and Pytorch baseline models, Paper: Lam et al. 2018
  - **Open AI Challenge: Aerial Imagery of South Pacific Islands** (*WeRobotics & Worldbank, May 2018*)  
Tree position & 4 tree species, RGB UAV imagery (0.4m/0.8m res.), multiple AOIs in Tonga
  - **NIST DSE Plant Identification with NEON Remote Sensing Data** (*inria.fr, Oct 2017*)  
Tree position, tree species and crown parameters, hyperspectral (1m res.) & RGB imagery (0.25m res.), LiDAR point cloud and canopy height model
  - **NOAA Fisheries Steller Sea Lion Population Count** (*NOAA, Jun 2017*)  
5 sea lion categories, ~ 80k instances, ~ 1k aerial images, Kaggle kernels
  - **Stanford Drone Data** (*Stanford University, Oct 2016*)  
60 aerial UAV videos over Stanford campus and bounding boxes, 6 classes (Pedestrian, Biker, Skateboarder, Cart, Car, Bus), Paper: Robicquet et al. 2016
  - **Cars Overhead With Context (COWC)** (*Lawrence Livermore National Laboratory, Sep 2016*)  
32k car bounding boxes, aerial imagery (0.15m res.), 6 cities, Paper: Mundhenk et al. 2016

### 3. Semantic Segmentation

- **FloodNet** (*University of Maryland, Jun 2021*)  
2343 image chips (drone imagery), 10 landcover categories (background, water, building flooded, building non-flooded, road-flooded, ...). Paper: Rahnemoonfar et al., 2021
- **LoveDA** (*Wuhan University, Oct 2021*)  
5987 image chips (Google Earth), 7 landcover categories, 166768 labels, 3 cities in China. Paper: Wang et al., 2021
- **FloodNet Challenge** (*UMBC, Microsoft, Texas A&M, Dewberry, May 2021*)  
2343 UAV images from after Hurricane Harvey, landcover labels (10 categories, e.g. building flooded, building non-flooded, road-flooded, ..), 2 competition tracks (Binary & semantic flood classification; Object counting & condition recognition)
- **Dynamic EarthNet Challenge** (*Planet, DLR, TUM, April 2021*)  
Weekly Planetscope time-series (3m res.) over 2 years, 75 aois, landcover labels (7 categories), 2 competition tracks (Binary land cover classification & multi-class change detection)

- 
- **Sentinel-2 Cloud Mask Catalogue** (Francis, A., et al., Nov 2020) 513 cropped subscenes (1022x1022 pixels) taken randomly from entire 2018 Sentinel-2 archive. All bands resampled to 20m, stored as numpy arrays. Includes clear, cloud and cloud-shadow classes. Also comes with binary classification tags for each subscene, describing what surface types, cloud types, etc. are present.
  - **MiniFrance** (Université Bretagne-Sud and ONERA, Jul 2020)  
2000 very high resolution aerial images over 16 cities in France (50cm res., from IGN BDORTHO), 16 landcover categories (Urban, Industrial, Pastures, Forests, etc. from Copernicus UrbanAtlas 2012), designed for *semi-supervised* semantic segmentation. Paper: Castillo-Navarro et al., 2021
  - **LandCoverNet: A Global Land Cover Classification Training Dataset** (Alemohammad S.H., et al., Jul 2020) Version 1.0 of the dataset that contains data across Africa, (20% of the global dataset). 1980 image chips of 256 x 256 pixels in V1.0 spanning 66 tiles of Sentinel-2. Classes: water, natural bare ground, artificial bare ground, woody vegetation, cultivated vegetation, (semi) natural vegetation, and permanent snow/ice. Citation: Alemohammad S.H., et al., 2020 and blog post
  - **LandCover.ai: Dataset for Automatic Mapping of Buildings, Woodlands and Water from Aerial Imagery** (Boguszewski, A., et al., May 2020) 41 orthophotos (9000x9000 px) over Poland, Aerial Imagery (25cm & 50cm res.), manual segmentations masks for Buildings, Woodland and Water, Paper: Boguszewski et al., 2020
  - **95-Cloud: A Cloud Segmentation Dataset** (S. Mohajerani et. al., Jan 2020)  
34701 manually segmented 384x384 patches with cloud masks, Landsat 8 imagery (R,G,B,NIR; 30 m res.), Paper: Mohajerani et al. 2021
  - **Open Cities AI Challenge** (GFDRR, Mar 2020) .  
790k building footprints from Openstreetmap (2 label quality categories), aerial imagery (0.03-0.2m resolution, RGB, 11k 1024x1024 chips, COG format), 10 cities in Africa.
  - **DroneDeploy Segmentation Dataset** (DroneDeploy, Dec 2019)  
Drone imagery (0.1m res., RGB), labels (7 land cover categories: building, clutter, vegetation, water, ground, car) & elevation data, baseline model implementation.
  - **SkyScapes: Urban infrastructure & lane markings** (DLR, Nov 2019)  
Highly accurate street lane markings (12 categories e.g. dash line, long line, zebra zone) & urban infrastructure (19 categories e.g. buildings, roads, vegetation). Aerial imagery (0.13 m res.) for 5.7 km2 of Munich, Germany. Paper: Azimi et al. 2019
  - **Open AI Challenge: Caribbean** (MathWorks, WeRobotics, Wordlbank, DrivenData, Dec 2019)  
Predict building roof type (5 categories, e.g. concrete, metal etc.) of provided building footprints

---

(22,553), RGB UAV imagery (4cm res., 7 areas in 3 Caribbean countries)

- **SpaceNet 5: Automated Road Network Extraction & Route Travel Time Estimation** (*CosmiQ Works, Maxar, Intel, AWS, Sep 2019*)  
2300 image chips, street geometries with location, shape and estimated travel time, 3/8band Worldview-3 imagery (0.3m res.), 4 global cities, 1 holdout city for leaderboard evaluation, APLS metric, baseline model
- **SEN12MS** (*TUM, Jun 2019*)  
180,748 corresponding image triplets containing Sentinel-1 (VV&VH), Sentinel-2 (all bands, cloud-free), and MODIS-derived land cover maps (IGBP, LCCS, 17 classes, 500m res.). All data upsampled to 10m res., georeferenced, covering all continents and meteorological seasons, Paper: Schmitt et al. 2018
- **Slovenia Land Cover Classification** (*Sinergise, Feb 2019*)  
10 land cover classes, temporal stack of hyperspectral Sentinel-2 imagery (R,G,B,NIR,SWIR1,SWIR2; 10 m res.) for year 2017 with cloud masks, Official Slovenian land use land cover layer as ground truth.
- **ALCD Reference Cloud Masks** (*CNES, Oct 2018*)  
8 classes (inc. cloud and cloud shadow) for 38 Sentinel-2 scenes (10 m res.). Manual labeling & active learning, Paper: Baetens et al. 2019
- **Agricultural Crop Cover Classification Challenge** (*CrowdANALYTIX, Jul 2018*)  
2 main categories corn and soybeans, Landsat 8 imagery (30m res.), USDA Cropland Data Layer as ground truth.
- **RoadNet** (*Wuhan, Oct 2018*)  
Road network labels, high-res Google Earth imagery, 21 regions, Paper: Liu et al. 2018
- **SpaceNet 3: Road Network Detection** (*CosmiQ Works, Radiant Solutions, Feb 2018*)  
8000 km of roads in 5 city aois, 3/8band Worldview-3 imagery (0.3m res.), SpaceNet Challenge Asset Library, Paper: Van Etten et al. 2018
- **Urban 3D Challenge** (*USSOCOM, Dec 2017*)  
157k building footprint masks, RGB orthophotos (0.5m res.), DSM/DTM, 3 cities, SpaceNet Challenge Asset Library
- **DSTL Satellite Imagery Feature Detection Challenge** (*Dstl, Feb 2017*)  
10 land cover categories from crops to vehicle small, 57 1x1km images, 3/16-band Worldview 3 imagery (0.3m-7.5m res.), Kaggle kernels
- **SPARCS: S2 Cloud Validation data** (*USGS, 2016*)  
7 categories (cloud, cloud shadows, cloud shadows over water, water etc.), 80 1kx1k px. subset Landsat 8 scenes (30m res.), Paper: Hughes, J.M. & Hayes D.J. 2014

- 
- **Biome: L8 Cloud Cover Validation data** (*USGS, 2016*)  
4 cloud categories (cloud, thin cloud, cloud shadows, clear), 96 Landsat 8 scenes (30m res.), 12 biomes with 8 scenes each, Paper: Foga et al. 2017
  - **Inria Aerial Image Labeling** (*inria.fr*)  
Building footprint masks, RGB aerial imagery (0.3m res.), 5 cities
  - **ISPRS Potsdam 2D Semantic Labeling Contest** (*ISPRS*)  
6 urban land cover classes, raster mask labels, 4-band RGB-IR aerial imagery (0.05m res.) & DSM, 38 image patches

#### 4. Scene classification

- **Airbus Wind Turbine Patches** (*Airbus, Mar 2021*)  
155k 128x128px image chips with wind turbines (SPOT, 1.5m res.).
- **BigEarthNet: Large-Scale Sentinel-2 Benchmark** (*TU Berlin, Jan 2019*)  
Multiple landcover labels per chip based on CORINE Land Cover (CLC) 2018, 590,326 chips from Sentinel-2 L2A scenes (125 Sentinel-2 tiles from 10 European countries, 2017/2018), 66 GB archive, Paper: Sumbul et al. 2019
- **WiDS Datathon 2019 : Detection of Oil Palm Plantations** (*Global WiDS Team & West Big Data Innovation Hub, Jan 2019*) Prediction of presence of oil palm plantations, Planet satellite imagery (3m res.), ca. 20k 256 x 256 pixel chips, 2 categories oil-palm and other, annotator confidence score.
- **So2Sat LCZ42** (*TUM Munich & DLR, Aug 2018*)  
Local climate zone classification, 17 categories (10 urban e.g. compact high-rise, 7 rural e.g. scattered trees), 400k 32x32 pixel chips covering 42 cities (LCZ42 dataset), Sentinel 1 & Sentinel 2 (both 10m res.), 51 GB
- **Cactus Aerial Photos** (*CONACYT Mexico, Jun 2018*)  
17k aerial photos, 13k cactus, 4k non-actus, Kaggle kernels, Paper: López-Jiménez et al. 2019
- **Statoil/C-CORE Iceberg Classifier Challenge** (*Statoil/C-CORE, Jan 2018*)  
2 categories ship and iceberg, 2-band HH/HV polarization SAR imagery, Kaggle kernels
- **Functional Map of the World Challenge** (*IARPA, Dec 2017*)  
63 categories from solar farms to shopping malls, 1 million chips, 4/8 band satellite imagery (0.3m res.), COCO data format, baseline models, Paper: Christie et al. 2017
- **EuroSAT** (*DFK, Aug 2017*)  
10 land cover categories from industrial to permanent crop, 27k 64x64 pixel chips, 3/16 band Sentinel-2 satellite imagery (10m res.), covering cities in 30 countries, Paper: Helber et al. 2017

- 
- **Planet: Understanding the Amazon from Space** (*Planet, Jul 2017*)  
13 land cover categories + 4 cloud condition categories, 4-band (RGB-NIR) satellite imagery (5m res.), Amazonian rainforest, Kaggle kernels
  - **AID: Aerial Scene Classification** (*Xia et al., 2017*)  
10000 aerial images within 30 categories (airport, bare land, baseball field, beach, bridge, ...) collected from Google Earth imagery. Paper: Xia et al. 2017
  - **RESISC45** (*Northwestern Polytechnical University NWPU, Mar 2017*)  
45 scene categories from airplane to wetland, 31,500 images (700 per category, 256x256 px), image chips taken from Google Earth (rich image variations in resolution, angle, geography all over the world), Download Link, Paper: Cheng et al. 2017
  - **DeepSAT: SAT-4/SAT-6 airborne datasets** (*Louisiana State University, 2015*)  
6 land cover categories, 400k 28x28 pixel chips, 4-band RGBNIR aerial imagery (1m res.) extracted from the 2009 National Agriculture Imagery Program (NAIP), Paper: Basu et al. 2015
  - **UC Merced Land Use Dataset** (*UC Merced, Oct 2010*)  
21 land cover categories from agricultural to parkinglot, 100 chips per class, aerial imagery (0.30m res.), Paper: Yang & Newsam 2010

## 5. Other Focus / Multiple Tasks

- **SEN12MS-CR & SEN12MS-CR-TS** (*TUM, Jun 2020*)  
A multi-modal and mono-temporal data set for cloud removal. Sentinel-1 & Sentinel-2, 2018. 175 globally distributed aois. Paper: SEN12MS-CR - Ebel et al. 2020, SEN12MS-CR-TS - Ebel et al. 2020
- **IEEE Data Fusion Contest 2022** (*IEEE GRSS, Université Bretagne-Sud, ONERA, ESA, Jan 2022*)  
Semi-supervised semantic segmentation, 19 cities and surroundings with multi-sensor tiles (VHR Aerial imagery 50cm res., Elevation model) & per pixel labels (contains landcover / landuse classes from UrbanAtlas 2012), Data
- **IEEE Data Fusion Contest 2021** (*IEEE, HP, SolarAid, Data Science Experts, Mar 2021*)  
Detection of settlements without electricity, 98 multi-temporal/multi-sensor tiles ( Sentinel-1, Sentinel-2, Landsat-8, VIIRS), per chip & per pixel labels (contains buildings, presence electricity).
- **University-1652: Drone-based Geolocalization (Image Retrieval)** (*ACM Multimedia, Oct 2020*)  
Corresponding imagery from drone, satellite and ground camera of 1,652 university buildings, Paper: Zheng et al. 2020



- 
- **IEEE Data Fusion Contest 2020** (*IEEE & TUM, Mar 2020*)  
Land cover classification based on SEN12MS dataset (see category Semantic Segmentation on this list), low- and high-resolution tracks.
  - **IEEE Data Fusion Contest 2019** (*IEEE, Mar 2019*)  
Multiple tracks: Semantic 3D reconstruction, Semantic Stereo, 3D-Point Cloud Classification. Worldview-3 (8-band, 0.35cm res.) satellite imagery, LiDAR (0.80m pulse spacing, ASCII format), semantic labels, urban setting USA, baseline methods provided, Paper: Le Saux et al. 2019 Outcome Part A: Kunwar et al. 2020 Outcome Part B: Lian et al. 2020
  - **IEEE Data Fusion Contest 2018** (*IEEE, Mar 2018*)  
20 land cover categories by fusing three data sources: Multispectral LiDAR, Hyperspectral (1m), RGB imagery (0.05m res.), Paper: Xu et al. 2019
  - **DEEPGLOBE - 2018 Satellite Challenge** (*CVPR, Apr 2018*)  
Three challenge tracks: Road Extraction, Building Detection, Land cover classification, Paper: Demir et al. 2018
  - **TiSeLaC: Time Series Land Cover Classification Challenge** (*UMR TETIS, Jul 2017*)  
Land cover time series classification (9 categories), Landsat-8 (23 images time series, 10 band features, 30m res.), Reunion island
  - **Multi-View Stereo 3D Mapping Challenge** (*IARPA, Nov 2016*)  
Develop a Multi-View Stereo (MVS) 3D mapping algorithm that can convert high-resolution Worldview-3 satellite images to 3D point clouds, 0.2m lidar ground truth data.
  - **Draper Satellite Image Chronology** (*Draper, Jun 2016*)  
Predict the chronological order of images taken at the same locations over 5 days, Kaggle kernels

## More Resources

- **awesome-remote-sensing-change-detection**
- **Radiant MLHub Training Data Registry**
- **satellite-image-deep-learning**