Formulating Machine Learning Problems for Satellite Imagery
(and solving them)

Valentina Staneva, Data Scientist, eScience Institute

- Discussion on different ML formulations of satellite imagery problems
- Overview of popular computer vision tasks and how they have been addressed with the recent deep learning advances
    - demystify some jargon
    - provide references to Python implementations
- Go through a couple of example notebooks
- Steps for building our own training datasets

Supervised Learning

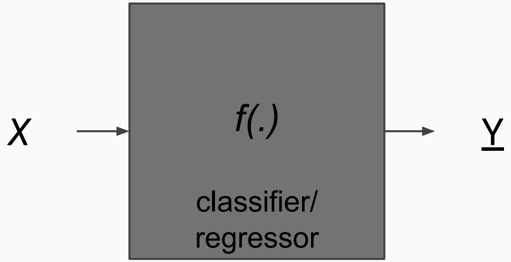
Labeled observations:

X_1, Y_1

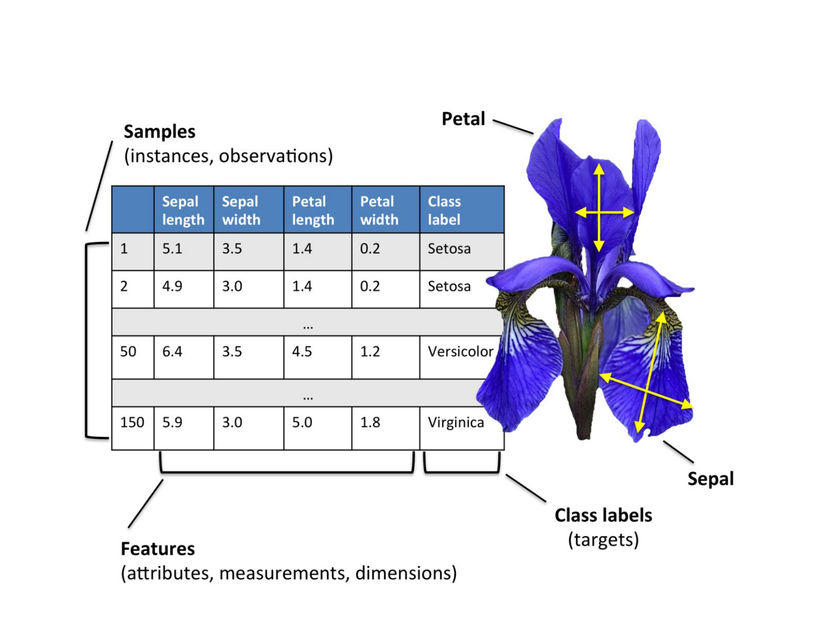
X_2, Y_2

…

X_n, Y_n



Find *f* to minimize the expected *loss(f(X),Y)*.



Source:
https://setscholarsanalytics.podia.com/applied-machine-learning-using-python-classification-with-iris-dataset

Supervised Learning for Satellite Imagery

What are X and Y?

- **Images:** Collection of images with corresponding labels
  - categories: rural/urban, cloud/no cloud, snow/cloud, etc.
  - Image pairs: other modalities, other sensors, other time stamps, other resolutions

- **Pixels:** One (or more) image with each pixel labeled to be a certain class (land cover segmentation)

- **Regions:** represented by bounding boxes, polygons, or masks
  - Labels for each pixel
  - Labels only for objects of interest

Supervised Learning for Satellite Imagery

Popular Computer Vision Tasks

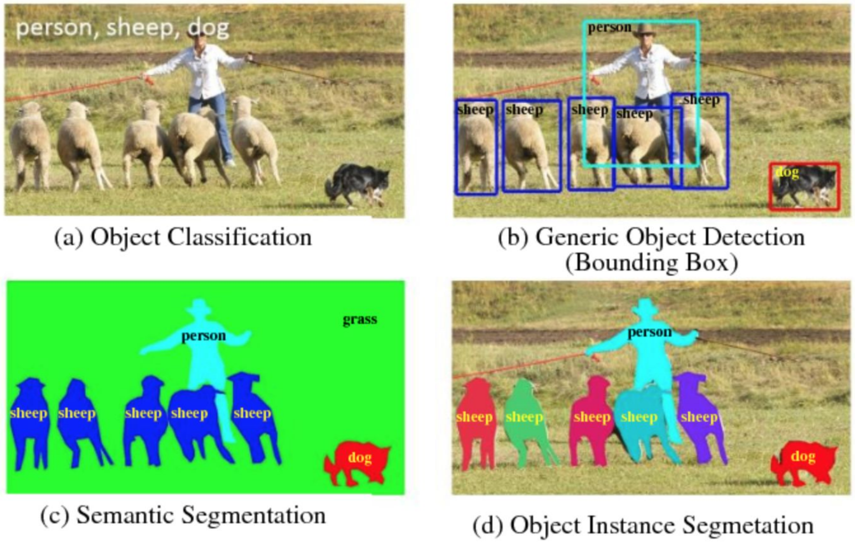
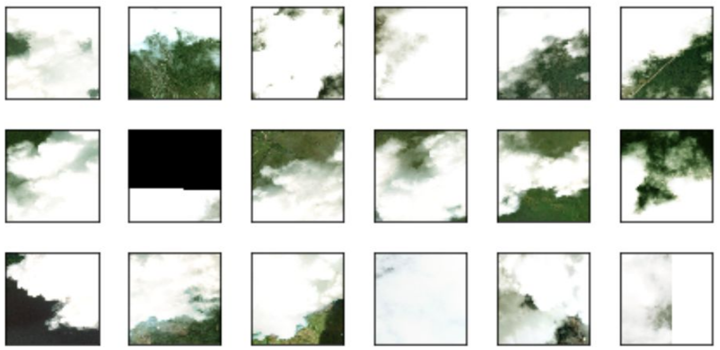


Fig. 3 Recognition problems related to generic object detection. (a) Image level object classification, (b) bounding box level generic object detection, (c) pixel-wise semantic segmentation, (d) instance level semantic segmentation.

Source: [Deep Learning for Generic Object Detection: A survey, Liu et al](#)

Image Classification



Approach 1:

- Extract features & apply scikit-learn classifiers

Approach 2:

- Use Convolutional Neural Networks on the full images
  - There exist pretrained models which can be used as a starting point
    - VGG, ResNet, MobileNet
  - Easy implementation in Keras
  - Image augmentation
    - Rotation
    - Translation
    - Affine transformation (nadir)
    - Resolution/blur
    - Light, color
    - Cloud occlusion

Evaluation:

- Standard classification metrics apply:
  - Accuracy, precision, recall, F1 score

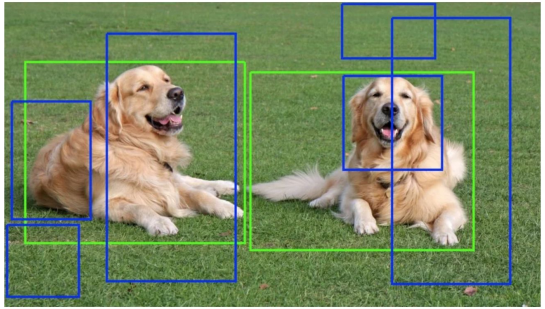
Object Detection

How can we identify more than one object in an image?

Approach 1:
Split into small windows and classify each window
- hopefully only one object falls into the small window
- too expensive to apply classification to each small window

Approach 2:
1) propose regions (using a small network)
2) classify only those regions, combine bounding boxes

- Faster RCNN (Region Conv. Neural Networks)
- Yolov3 (You Only Look Once)
- SSD (Single Shot Detector)


Blue Boxes: False Positives; Green Boxes: True Positives

https://github.com/tensorflow/models/tree/master/research/object_detection
https://github.com/qqwweee/keras-yolo3

Evaluation: mean Average Precision (mAP)

Semantic Segmentation

X_i - image, Y_i - mask

Approach 1:

- Pixel by pixel classification (assume independence)

Approach 2:

- Fully convolutional neural networks (keep pixels together)
- U-net (developed for biomedical imaging, works with little training data)
- Easy Implementation in keras

Evaluation:

- Evaluation based on standard metrics (like pixel accuracy, F1 scores): no geometric information
- Intersection/total area, boundary F1 scores

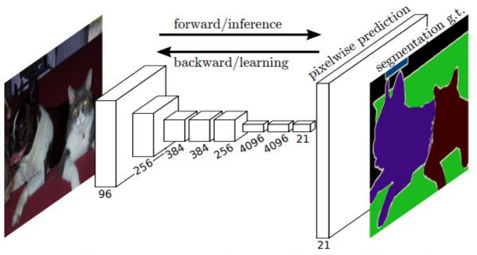
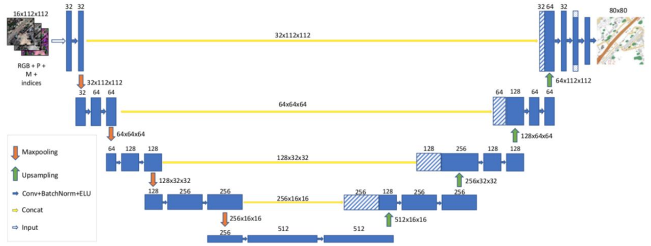


Figure 1. Fully convolutional networks can efficiently learn to make dense predictions for per-pixel tasks like semantic segmentation.

Source: [Fully Convolutional Networks for Semantic Segmentation, Long et. al.]



Source: [U-net architecture]

Mask RCNN

X_i - image, Y_i - mask

Approach 1:

- Segment directly

Approach 2:

- Detect bounding box
- Segment within bounding box

https://github.com/matterport/Mask_RCNN

Model predicting mask segmentations and bounding boxes for ships in a satellite image

Source:https://towardsdatascience.com/mask-r-cnn-for-ship-detection-segmentation-a1108b5a083

Example Notebooks

[Kaggle Airbus Ship Detection Challenge]

- Segmentation with U-net

https://www.kaggle.com/valcoder/ship-detection-using-keras-u-net

- Segmentation with Mask RCNN

https://www.kaggle.com/valcoder/mask-r-cnn-ship-detection-minimum-viable-model-1

Model predicting mask segmentations and bounding boxes for ships in a satellite image

Source:https://towardsdatascience.com/mask-r-cnn-for-ship-detection-segmentation-a1108b5a083

Data Preparation

Getting data into the right format for Machine Learning:

- Chipping images and preserving labels
- Geospatial coordinates ⇔ pixel coordinates
- Mask ⇔ boundary coordinates ⇔ geospatial coordinates
- Overlapping geospatial regions on geotiff images
- Merging chips for mapping

Tools:

- rasterio, geopandas, shapely
- robosat

Special Formats:

- COCO (Common Objects in Context) format
- Slippy Map Format (from OSM)
- Run Length Encoding for storing masks