Building Detection with the SpaceNet Dataset

6th semester project
Geospatial Big Data Analytics
National Technical University of Athens

Pavlos Kalfantis - MSc student

Part 1
Data Exploration

SpaceNet Challenges

- 8 SpaceNet mapping Challenges
- Each dataset focuses on a different aspect of ML to solve mapping challenges
  - Building detection
  - Road network detection
  - Road network extraction and travel time estimation
  - Multi sensor all-weather mapping
  - Multi temporal urban development
  - Flood detection
- This project: SpaceNet1 Dataset for **building detection**

SpaceNet1 Dataset for Building Detection

- Building detection is important in Urban Planning tasks
- Dataset with TorchGeo through Radiant MLHub (cloud based library)
- Dataset contains building footprints over the city of Rio de Janeiro
- 6940 8-band images - 6940 RGB images
- 382,534 polygons (labels) in **GeoJSON** format
- 8-band image - $1m^2$ x $1m^2$
- **RGB image used - $50cm^2$ x $50cm^2$ - 406x429 pixels (lower cost, effective)**

The data

RGB Image

GeoJSON label


```
{'crs': {'properties': {'name': 'urn:ogc:def:crs:OGC:1.3:CRS84'},
         'type': 'name'},
 'features': [{'geometry': {'coordinates': [[[-43.69487929999997,
                                                -22.984465499999942,
                                                0.0],
                                               [-43.69493559999995,
                                                -22.984406699999965,
                                                0.0],
                                               [-43.694879399999934,
                                                -22.984365499999967,
                                                0.0],
                                               [-43.694828,
                                                -22.98442369999998,
                                                0.0],
                                               [-43.69487929999997,
                                                -22.984465499999942,
                                                0.0]]],
                             'type': 'Polygon'},
               'properties': {'HGISOID': 946847.0,
                              'HGIS_OID': '946847.0',
                              'QAStatus': 'Original_Building',
                              'Revision1': 'No',
                              'Shape_Area': 0.0,
                              'Shape_Leng': 0.000295,
                              'TaskArea': 'West',
                              'area': 'None',
                              'building': 'yes',
                              'changeset': '5404',
                              'id': 'way/70607',
                              'partialBuilding': 0.0,
                              'partialDec': 1.0,
                              'timestamp': '2016-06-22T21:22:31Z',
                              'type': 'None',
```


Semantic Segmentation needs masks as target!

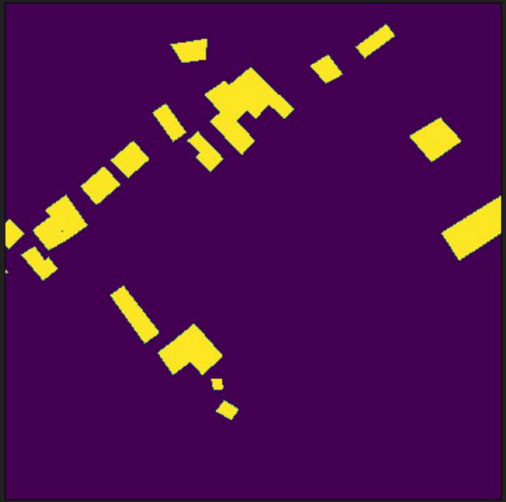
Ground Truth Transform

- Polygon vertices (latitude and longitude) to pixel coordinates -> Polygons
- Filled polygons -> Building masks

Build Torch Dataset

- Select **subset** of images for faster training
- 1000 images used, containing at least one building
- Crop 400x400 pixels of each image - standard size for Unet
- 70-20-10 Train - Validation - Test split
- Normalize images with mean and standard deviation of training set





Part 2
Literature Review

Approaches for building detection

- Semantic Segmentation
  - Fully Convolutional Architecture - SegNet
  - **UNet Architecture**
  - UNet with Residual blocks
  - UNet with Attention mechanism
- Object Detection
  - Faster RCNN
  - HOG Detector
  - YOLO - YOLT - **YOLT2**
- Conditional Random Fields

SpaceNet Challenge Metric

- Based on the Jaccard Index (IoU)
- Generated polygons to represent footprints
- Each proposed footprint is either a TP or FP
- TP if the proposal is closest to a labeled polygon and IoU>0.5
- After detection is defined, calculate precision, recall and **F1 score** (over ALL of the test imagery)
- Winning F1 score was 0.26!

Image Segmentation Metrics - pixelwise (simplified)

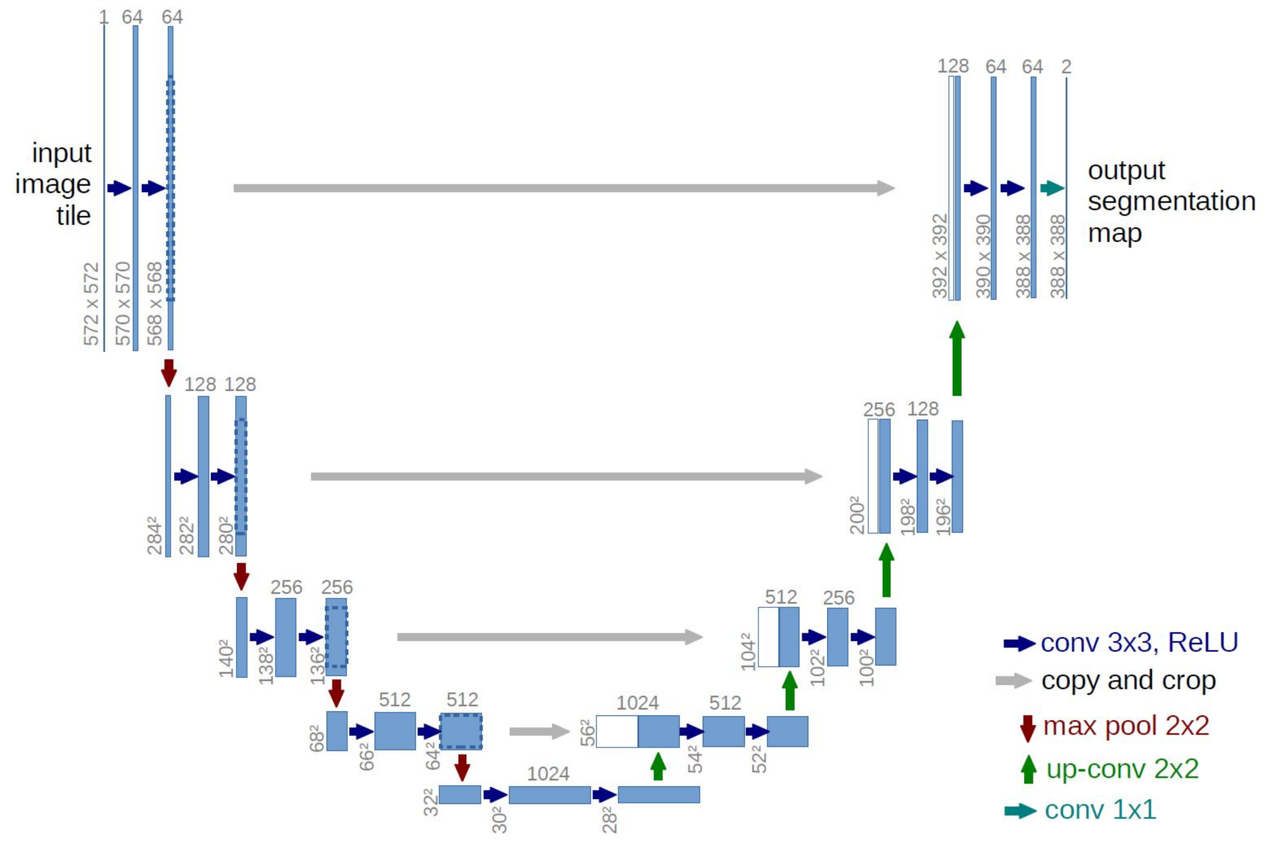
- Accuracy
- Precision
- Recall
- F1-Score
- IoU (Jaccard)

Architecture	Evaluation Metrics				
	Accuracy	Precision	Recall	F1	Jaccard
SegNet [27]	0.919	0.569	**0.813**	0.662	-
SegNet with Sobel filters [10]	0.923	0.596	0.722	0.667	-
CRF with Sobel filters [10]	0.931	0.632	0.763	0.675	-
CRF with CNN boundaries [10]	0.924	0.624	0.764	0.674	-
U-Net [29]	0.923	0.808	0.808	0.798	0.700
ResU-Net	0.936	**0.864**	0.770	0.811	0.703
Attention U-Net	**0.940**	0.851	0.809	**0.826**	**0.726**

Part 3
UNet Architecture for Semantic Segmentation

UNet Architecture

- Skip connections
- Symmetric
- 400x400 input



Part 4
Model Training and Metrics

Model Training Notes

- Complex loss function (dice loss + binary cross entropy)
- Early stopping on pixel accuracy
- Model trained for 18 epochs on Google Colab’s GPUs

(Binary) Cross Entropy (BCE): $$BCE = -\frac{1}{N}\sum_{i=0}^{N} y_i \cdot log(\hat{y}_i) + (1 - y_i) \cdot log(1 - \hat{y}_i)$$

$$soft\ Dice\ loss = 1 - \frac{2\sum_{pixels} y_{true} y_{pred}}{\sum_{pixels} y_{true}^2 + \sum_{pixels} y_{pred}^2}$$

Model Evaluation

- Results on test set (100 images of 400x400 pixels each)

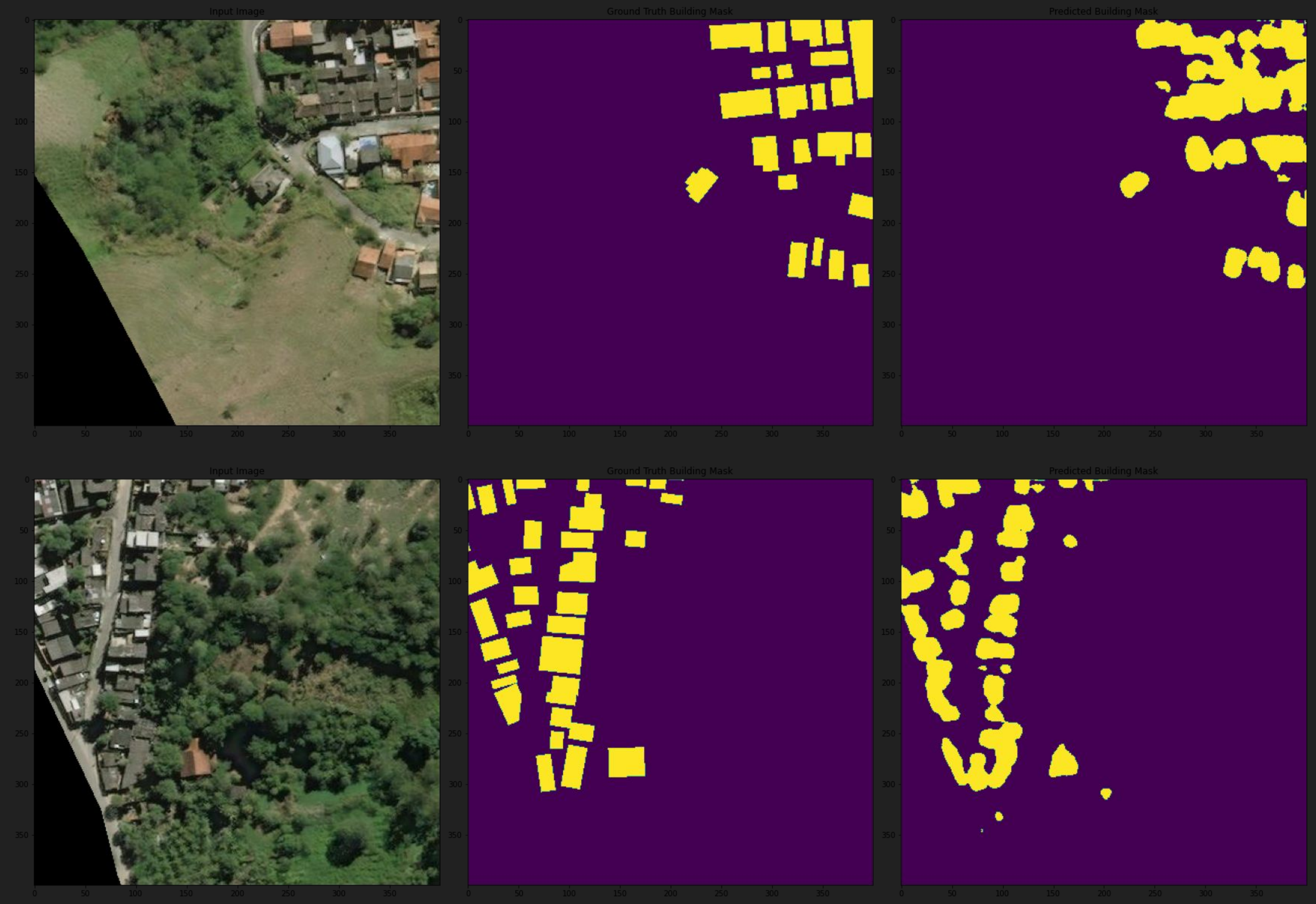
Model Results

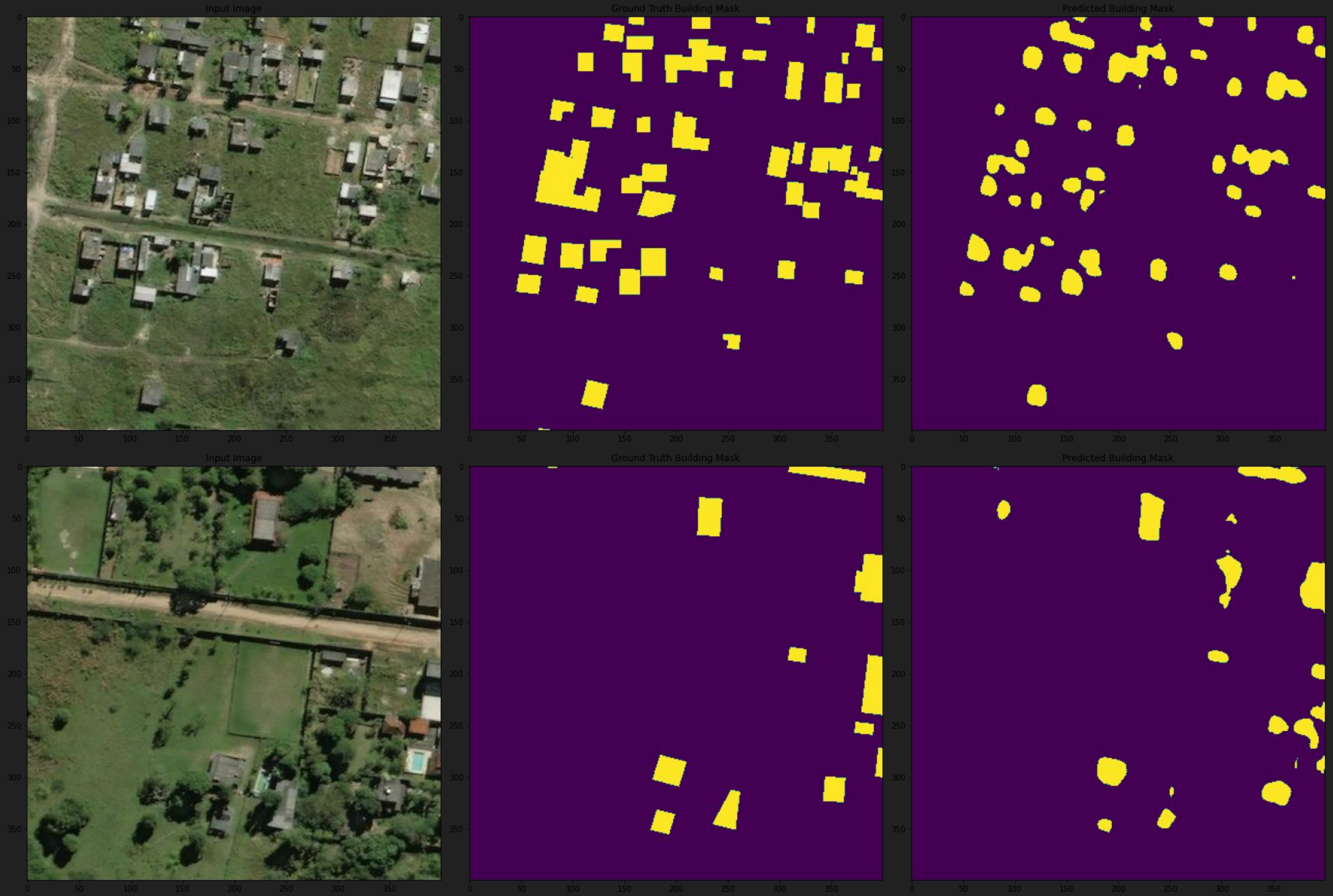
```
Accuracy:  0.9235985
Precision:  0.7395110456998663
Recall:  0.700275604734716
F1-score:  0.7193587252377506
IoU: 0.7385114431381226
```

Literature Results

Architecture	Evaluation Metrics				
	Accuracy	Precision	Recall	F1	Jaccard
SegNet [27]	0.919	0.569	**0.813**	0.662	-
SegNet with Sobel filters [10]	0.923	0.596	0.722	0.667	-
CRF with Sobel filters [10]	0.931	0.632	0.763	0.675	-
CRF with CNN boundaries [10]	0.924	0.624	0.764	0.674	-
U-Net [29]	0.923	0.808	0.808	0.798	0.700
ResU-Net	0.936	**0.864**	0.770	0.811	0.703
Attention U-Net	**0.940**	0.851	0.809	**0.826**	**0.726**

Part 5
Visualize Results


Input Image
Ground Truth Building Mask
Predicted Building Mask
Input Image
Ground Truth Building Mask
Predicted Building Mask
0 50 100 150 200 250 300 350


Input Image
Ground Truth Building Mask
Predicted Building Mask
Input Image
Ground Truth Building Mask
Predicted Building Mask

References

- https://spacenet.ai/spacenet-buildings-dataset-v1/
- https://www.mdpi.com/2227-7080/10/1/19
- https://medium.com/the-downlinq/getting-started-with-spacenet-data-827fd2ec9f53
- https://arxiv.org/pdf/1807.01232.pdf
- https://medium.com/the-downlinq/building-extraction-with-yolt2-and-spacenet-data-a926f9ffac4f
- https://arxiv.org/pdf/1505.04597.pdf