AI Black Belt

kaggle

DEEPER

Data Analytics

Classification in GeoSpatial Data Analysis

by Loïc Quertenmont, PhD

04/06/2019

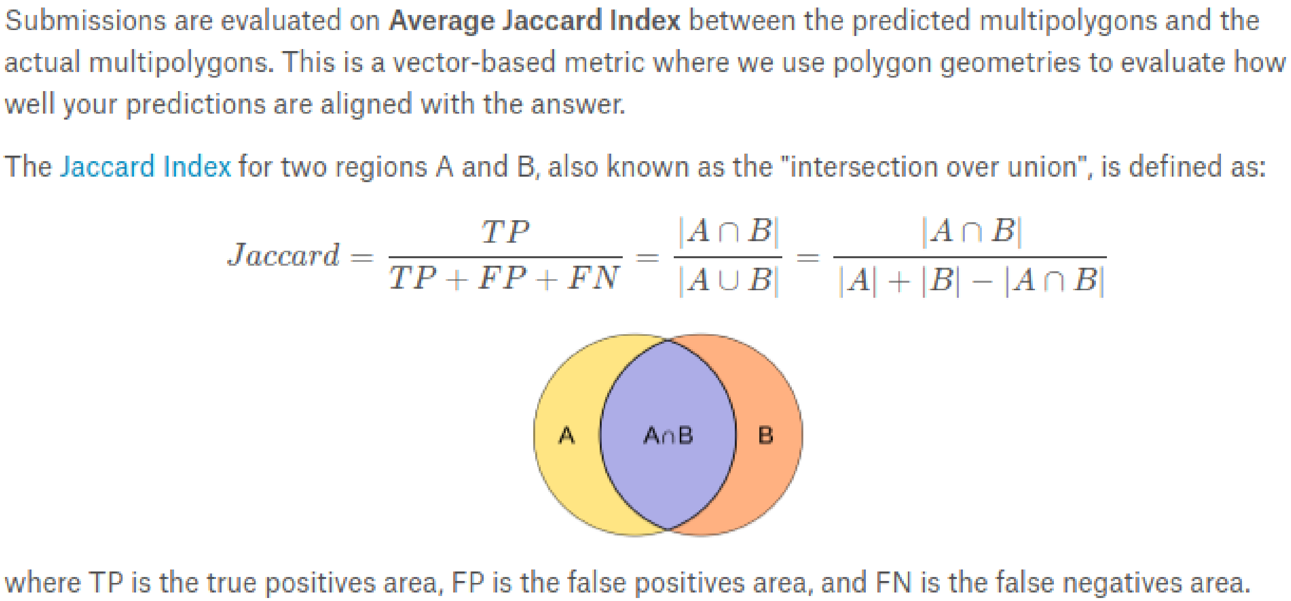
Yellow Belt

Context

- Kaggle competition:
  - Defence Science and Technology Laboratory (Dstl)
  - December 2016
  - Prize : $ 50K , $ 30K, $ 20K

  - Kagglers are challenged to **accurately classify features in overhead imagery**. Automating feature labeling will not only help Dstl make smart decisions more quickly around the defense and security of the UK, but also bring innovation to computer vision methodologies applied to satellite imagery.

  - Evaluation:



Submissions are evaluated on **Average Jaccard Index** between the predicted multipolygons and the actual multipolygons. This is a vector-based metric where we use polygon geometries to evaluate how well your predictions are aligned with the answer.

The Jaccard Index for two regions A and B, also known as the "intersection over union", is defined as:

$$Jaccard = \frac{TP}{TP + FP + FN} = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

where TP is the true positives area, FP is the false positives area, and FN is the false negatives area.

Context

- My goal :
  - Blog with full details : https://deeperanalytics.be/blog/
  - Demonstrate usage of big data tools with scalability in mind
    - Hadoop
    - Spark
    - H2O (Sparkling Water)
    - Geotrellis
  - ~~Win the competition~~
    - ~~Python~~
    - ~~State of the art Deep Learning techniques~~

Dataset

- DSTL provides **1km x 1km** satellite images in both 3-band and 16-band GeoTIFF formats. The images are coming from the WorldView 3 satellite sensor. In total, there are 450 images of which 25 have training labels → 45GB of data

- **R (3):** The RGB natural color images with an intensity resolution of 11-bits/pixel and a spatial resolution of **0.31m**.

- **P (1):** The 1 Panchromatic band (450-800 nm) has an intensity resolution of 11-bits/pixel and a spatial resolution of **0.31m**.

- **M (8):** The 8 Multispectral bands from 400 nm to 1040 nm (red, red edge, coastal, blue, green, yellow, near-IR1 and near-IR2) has an intensity resolution of 11-bits/pixel and a spatial resolution of **1.24m**.

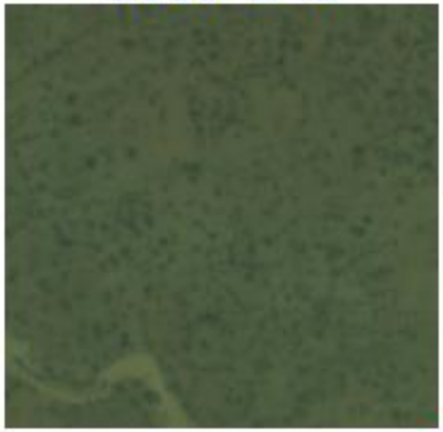
- **A (8):** The 8 short-wave infrared (SWIR) bands (1195 – 2365 nm) has an intensity resolution of 14-bits/pixel and a spatial resolution of **7.5m**

Dataset

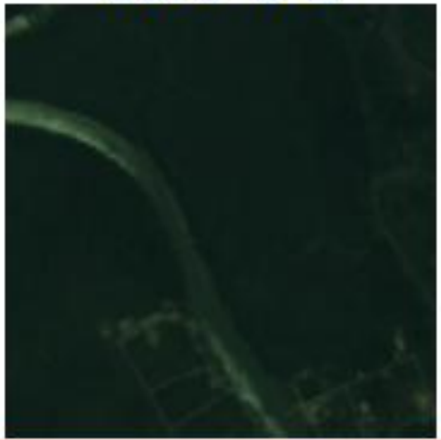
6040 2 2

6040 4 4

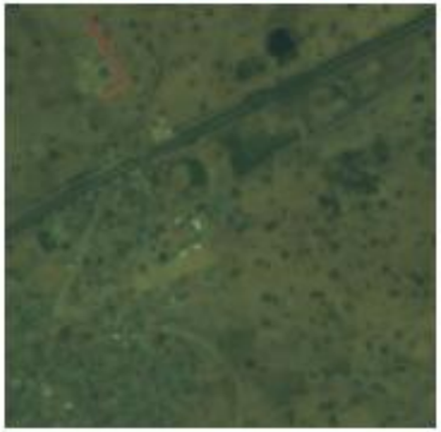
6060 2 3



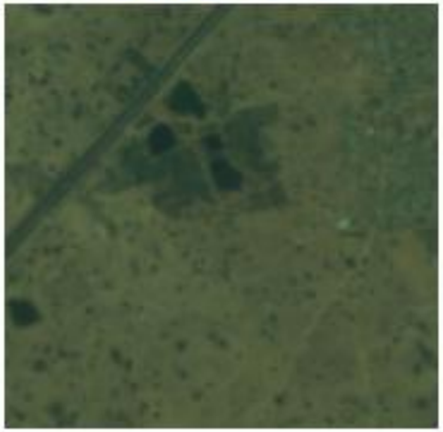
6070 2 3



6110 3 1

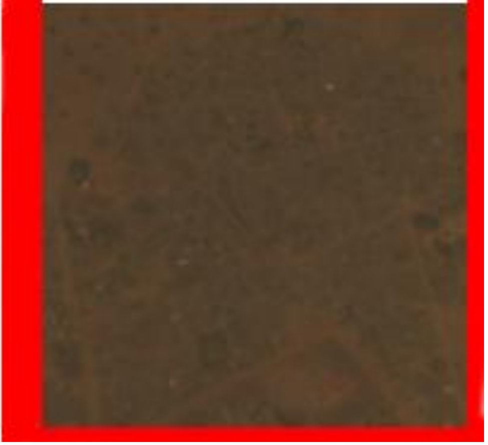


6110 4 0

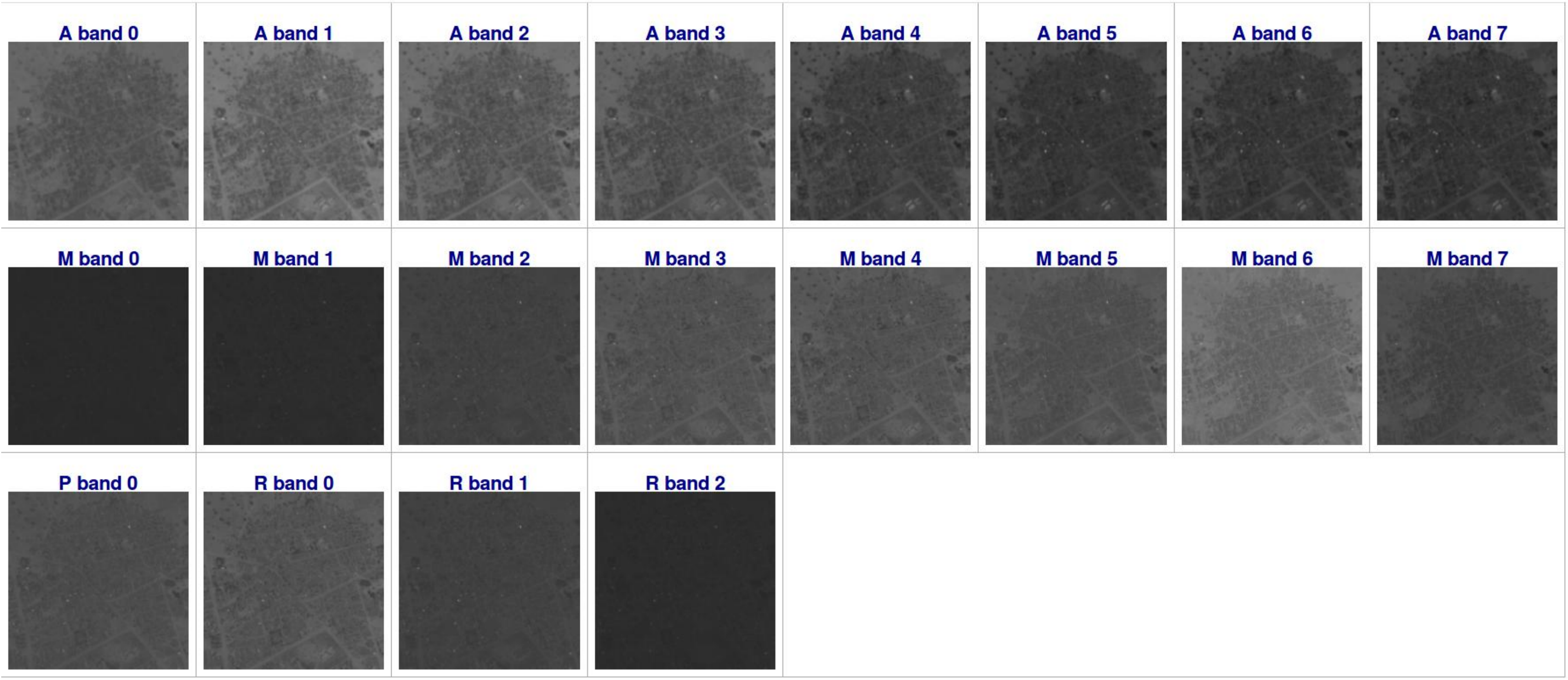


6120 2 0

6120 2 2



Dataset

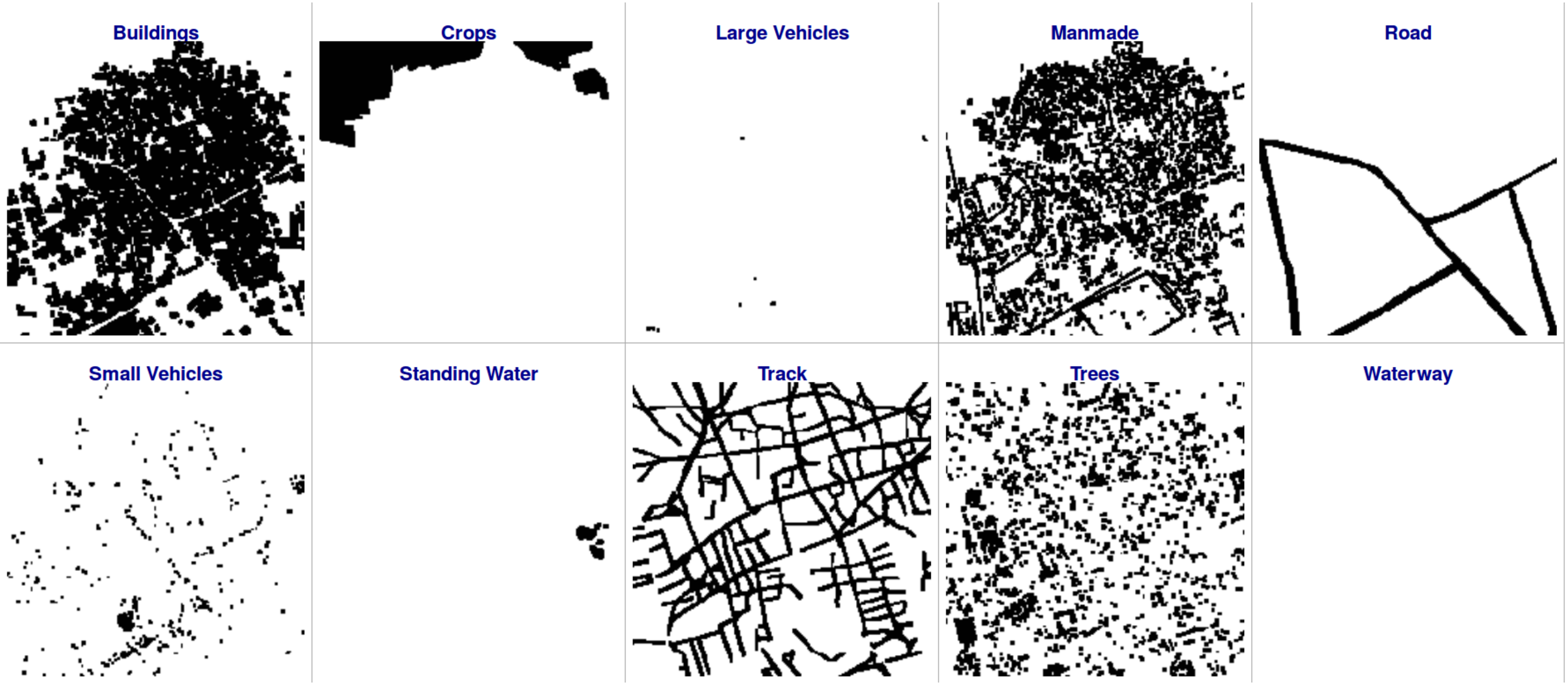

A band 0
A band 1
A band 2
A band 3
A band 4
A band 5
A band 6
A band 7
M band 0
M band 1
M band 2
M band 3
M band 4
M band 5
M band 6
M band 7
P band 0
R band 0
R band 1
R band 2

Dataset

- GeoJSON files containing identified (multi-)polygons on the 25 training images.
- There are polygons of each of the 10 possible object class types used in this competition:
    - **Buildings**
    - **Misc. Manmade structures**
    - **Road**
    - **Track**
    - **Trees**
    - **Crops**
    - **Waterway**
    - **Standing water**
    - **Large Vehicle**
    - **Small Vehicle**

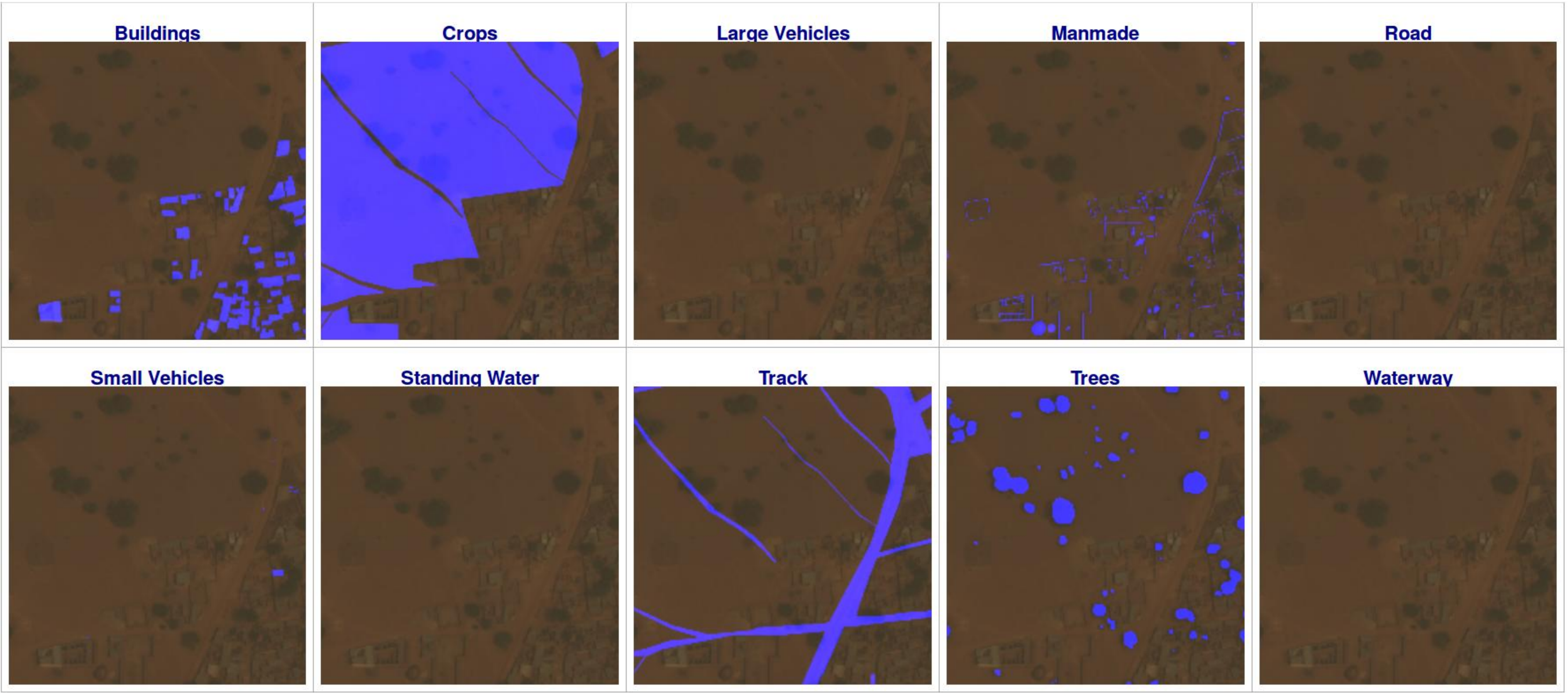
- NOTE: there are overlaps between classes
    - vehicles could be on Road
    - Trees in crops

Dataset


Buildings
Crops
Large Vehicles
Manmade
Road
Small Vehicles
Standing Water
Track
Trees
Waterway

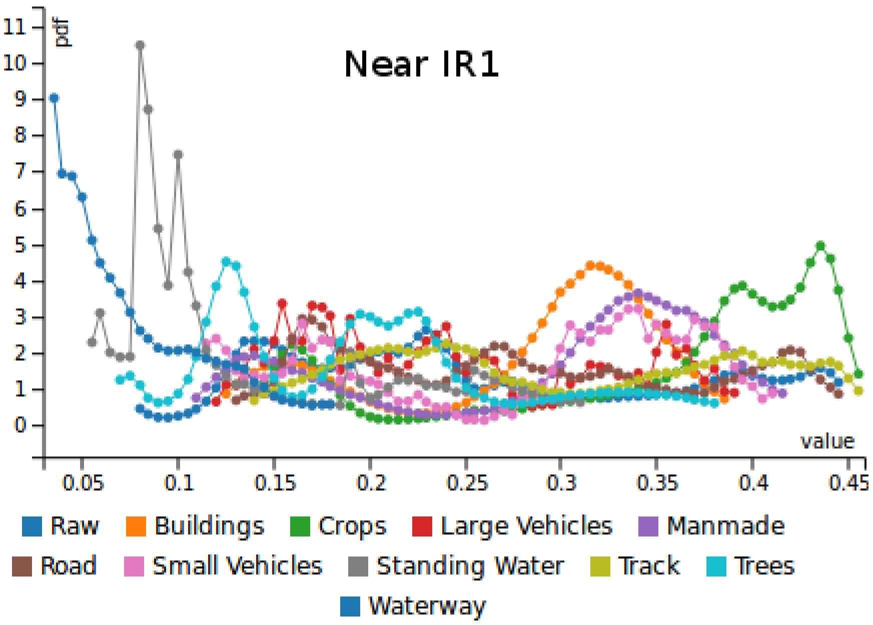
DEEPER
Data Analytics

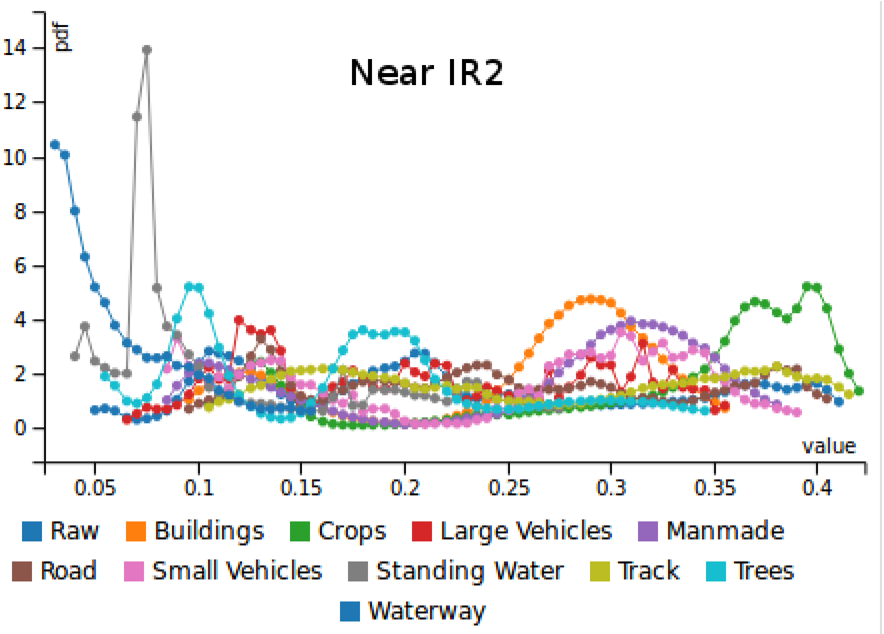
Combine & Zoom In



Preliminary Analysis

- Scale pixel intensity to be in the range [0,1] for all channels (according to their resolut.)
- Build pixel datasets in the training sample according to the class they belong to
- Check distribution of pixels according to their class for each spectral frequency

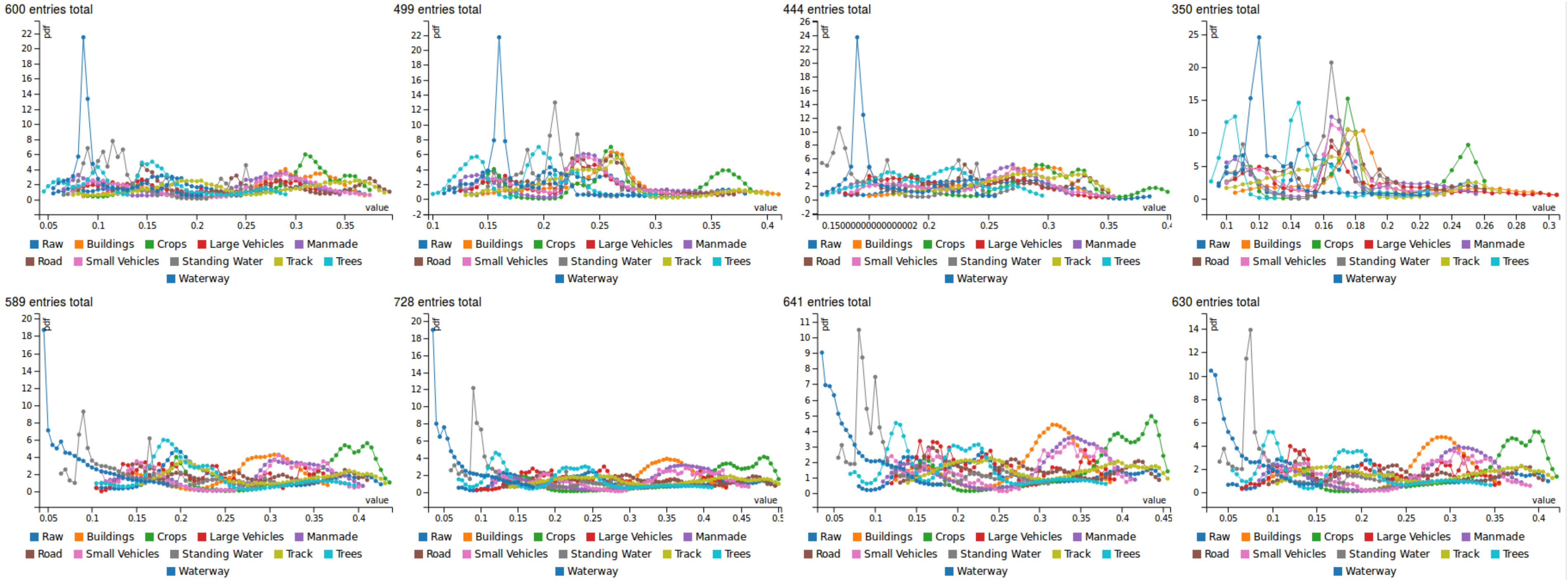




near-IR2 band is particularly good at discriminating water, building, and crops from the rest. Other bands might be more performent for other objects.

Preliminary Analysis

M bands:



Machine Learning

- **Classification problem… BUT**
  - One pixel can belong to more than one class
  - One vs All approach
    - 10 classes → 10 models to train → 10 probabilities of belonging to a given class
- **Simple approach**
  - Purely pixel based algorithm considering only one pixel data
    - Totally ignoring neighbor information at this stage
  - **For each class, we**
    - **Take 200K (random) pixels belonging to the class**
    - **Take 200K (random) pixels not belonging to the class**
    - **8+8+3+1 features for each pixel**
    - **1 boolean target : belong to class or NOT belong to class**
    - **Split: Train (90%) / Test (10%)**
    - **Train: a GBM (= super decision tree)**

DEEPER
Data Analytics

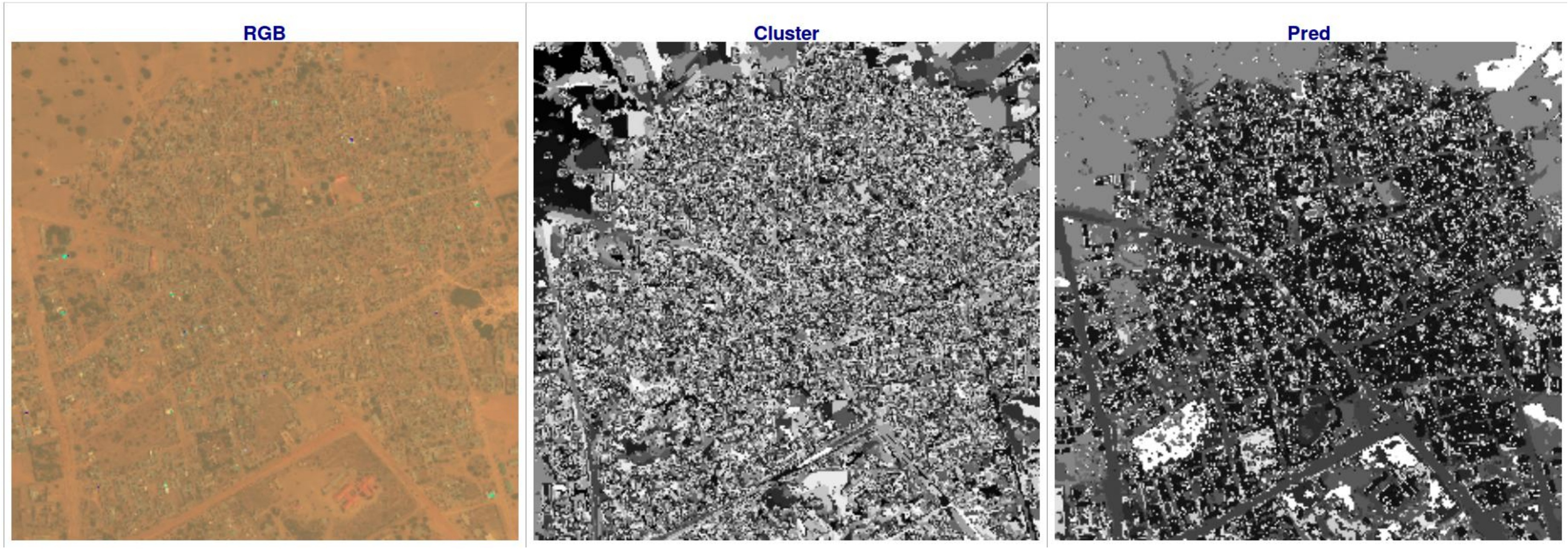
Machine Learning

Object Type	Model AUC
Buildings	0.992441
Manmade	0.966026
Road	0.997235
Track	0.930540
Trees	0.960790
Crops	0.983181
Waterway	0.999718
Standing Water	0.999889
Large Vehicles	0.999354
Small Vehicles	0.997031

Predict with the model

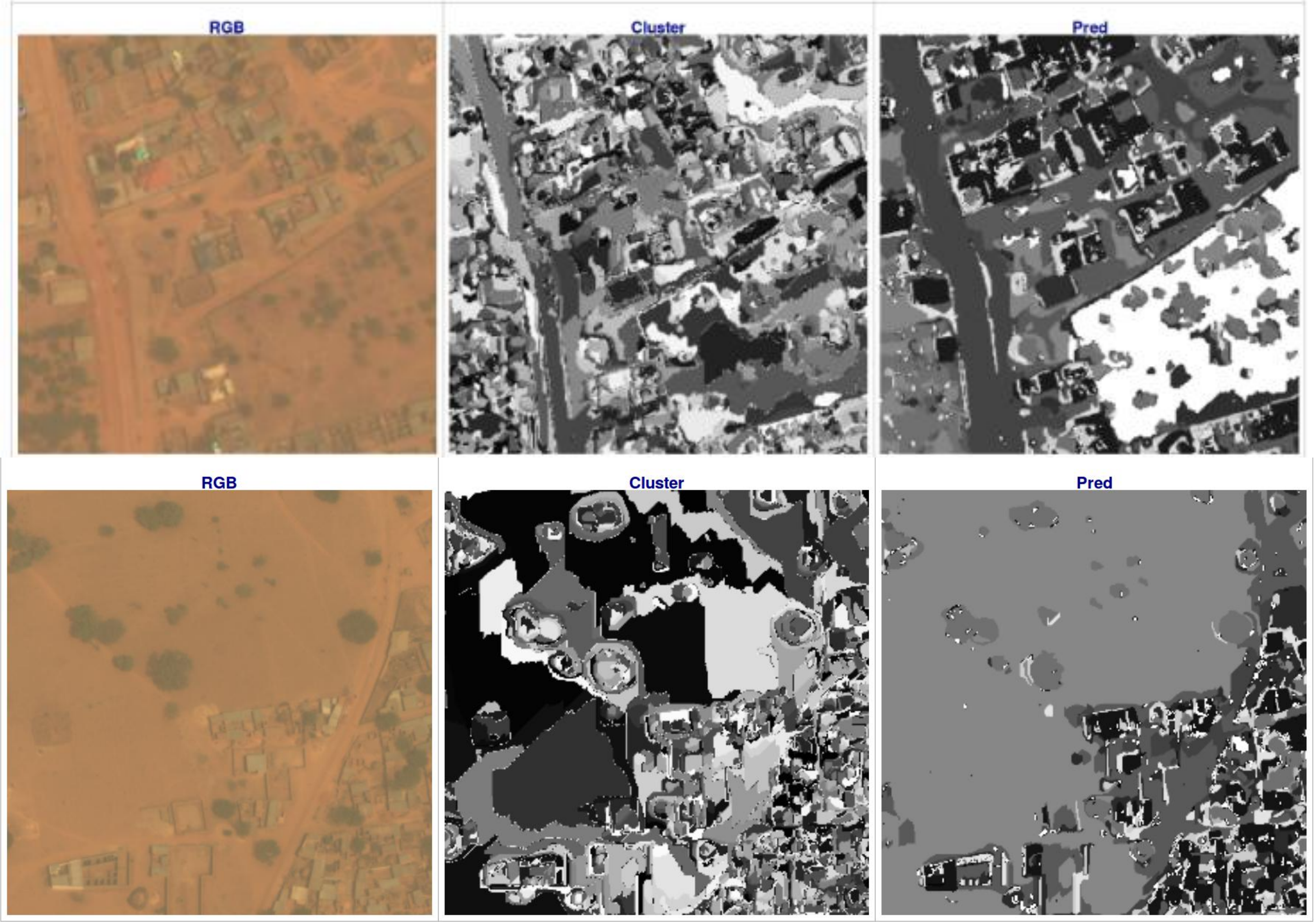
- Each of the 450 images have 11.5M pixels
  - We need to perform 10 classes * 450 images * 11.5M pixels predictions
  - =51Billions… This is sloooooowww…

- But… many adjacent pixels have the same spectral information (within 3%)
  - We can cluster similar (and adjacent) pixels in clusters
  - Run the model prediction only once per cluster (instead of once per pixel)



DEEPER

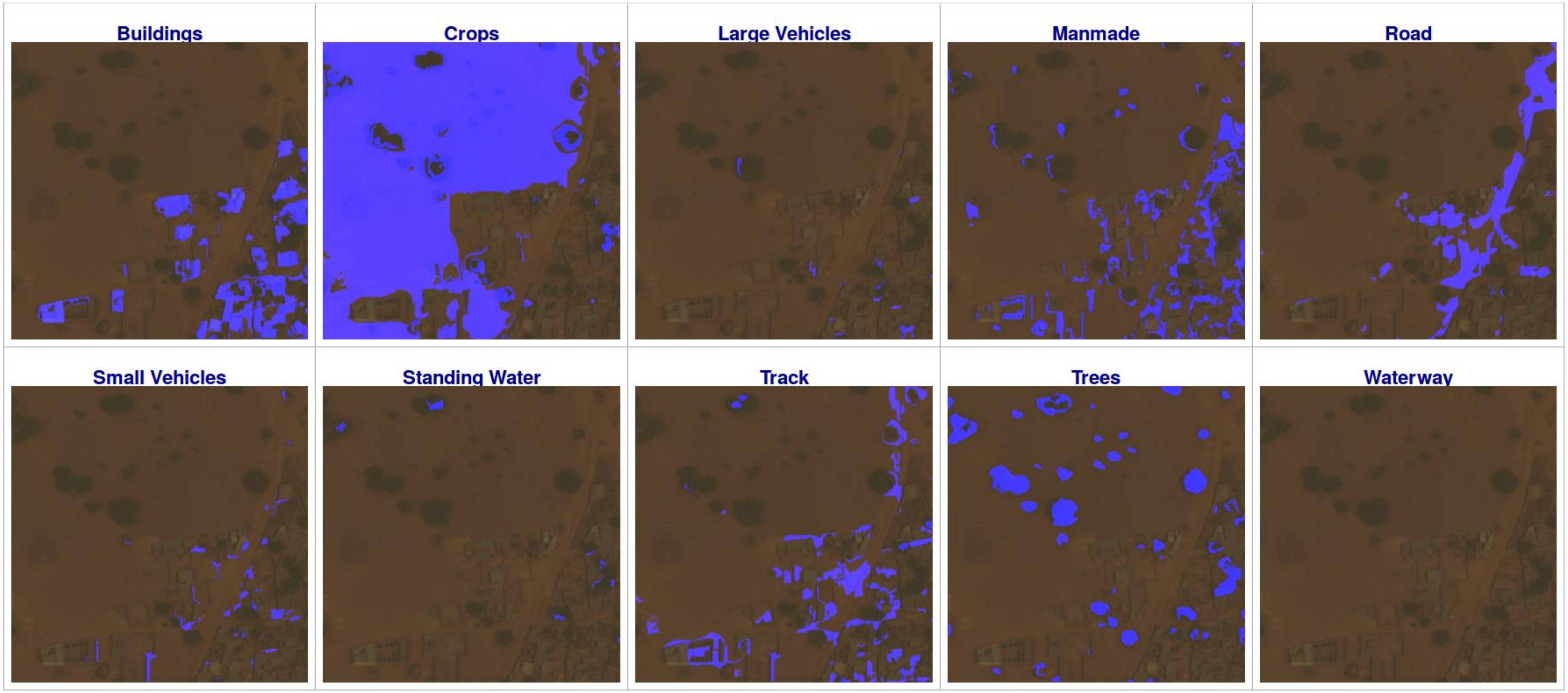
Predict with the model



DEEPER
Data Analytics

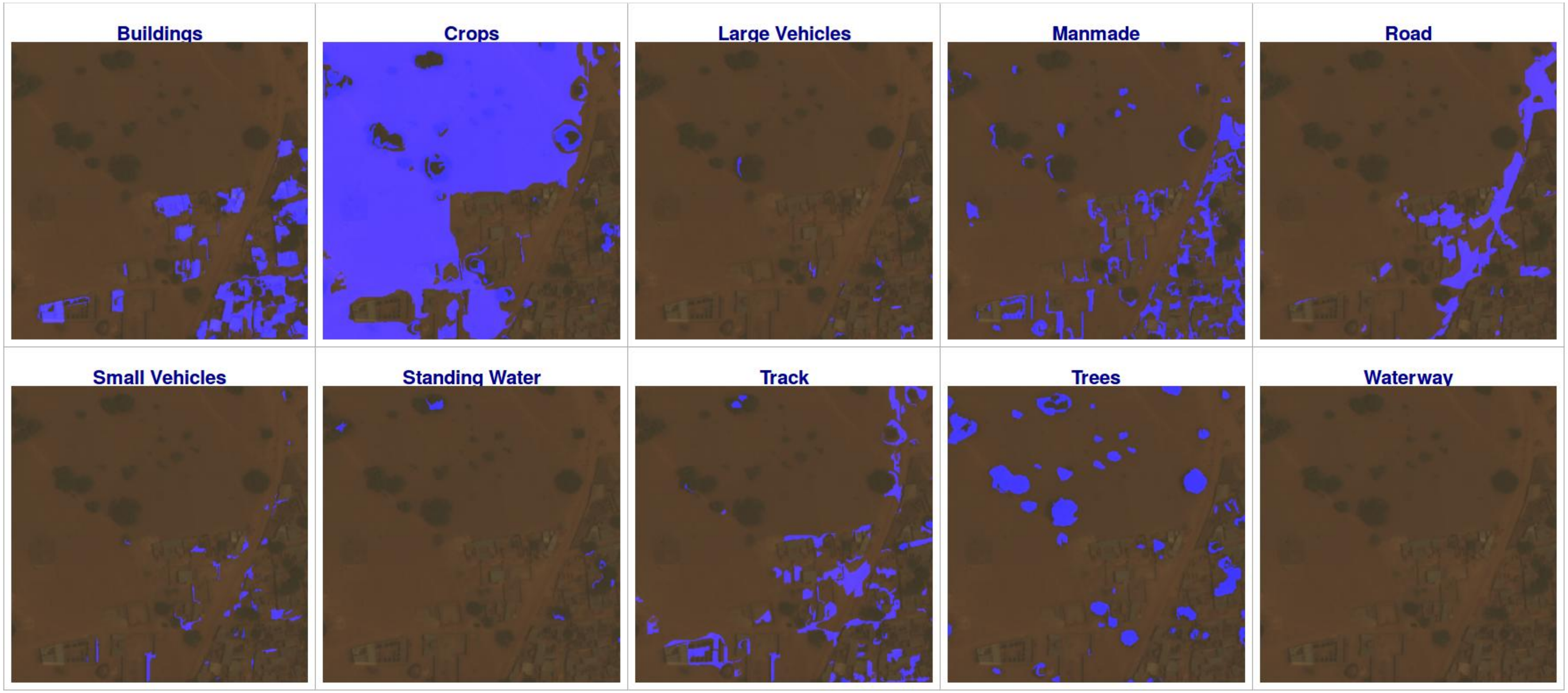
Building masks

- Flag classes of every cluster based on
    - Cluster and Neighboring cluster probabilities and
    - Ad-Hoc Rules
        - finding a truck in the middle of a waterway area are null
        - Finding tree on top of a road is unlikely
        - finding a car or a truck on a road is likely,
        - finding a tree in the middle of a crop is likely



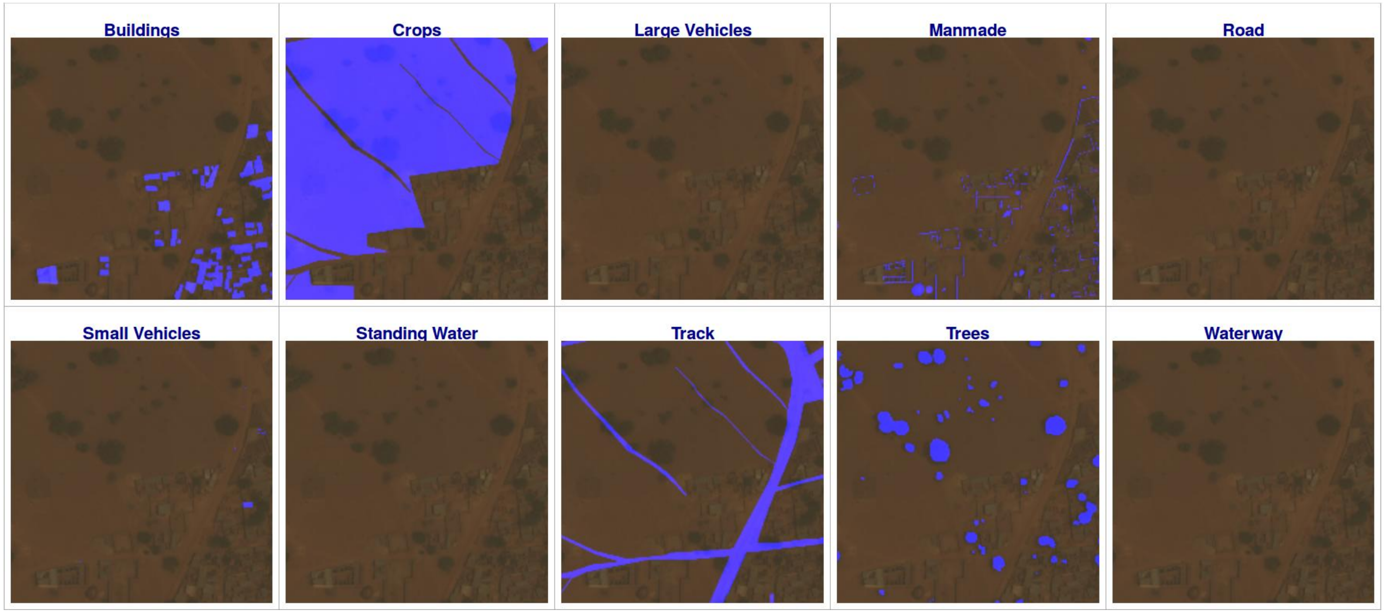
Building masks

- Flag classes of every cluster based on
  - Cluster and Neighboring cluster probabilities and
  - Ad-Hoc Rules
    - finding a truck in the middle of a waterway area are null
    - Finding tree on top of a road is unlikely
    - finding a car or a truck on a road is likely,
    - finding a tree in the middle of a crop is likely

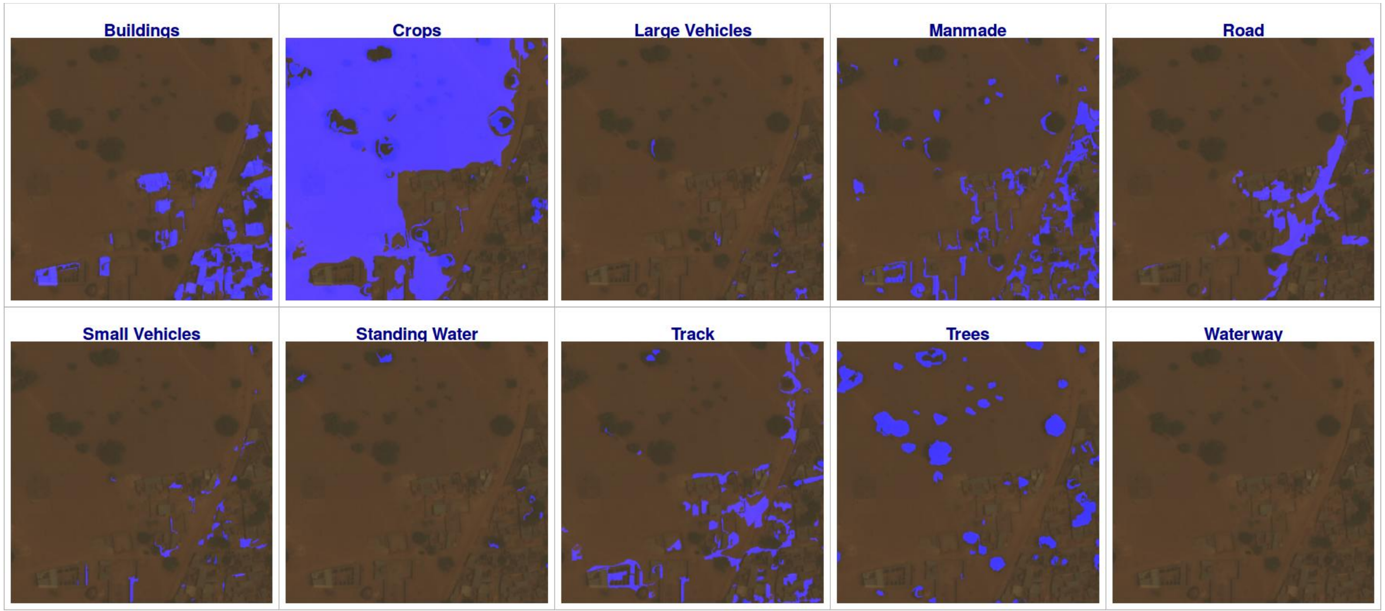


Results

Human Labeled



Our classification Model



Conclusion

- **That was my first time playing with satellite images**
- **Highly scalable → Every pixel/cluster can be ran independently**
- **Very Simple algorithm**
  - **Could do much better with state of the art tools**
    - **CNN based deep-learning (U-Net architecture)**
- **136th out of 419 competitors at Kaggle competition**

We do many other fun things at Deeper Analytics,
but also some serious business under NDA ;-)