Semantic Representations on Sattellite Images

Amy Loutfi
Center for Applied Autonomous Sensor Systems

Project: Large Scale Semantic Perception on Sattelite Imagery

Goal: Automatically augment sattelite imagery with annotations about objects, functions, and relations to enable 1) users to query the map 2) improve the recognition of objects and entities

Data

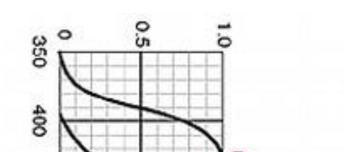
- 3D maps:
  - Formats: Ortho + DSM (GeoTIFF) and Collada.
  - Boden, Stockholm, part of Örebro (Collada only).
  - Multiple color bands to extract different features.
  - Resolution: 0.5m/px in all dimensions.
  - UTM coordinates, WGS84, ellipsoid.

Ortho + DSM

- Multiband + Synthetic Bands
- Ortho – aerial images are combined to remove projection tilts and displacements.
- DSM – surface model – per-pixel mapped to Ortho, elevation includes buildings and trees.
- GeoTiff

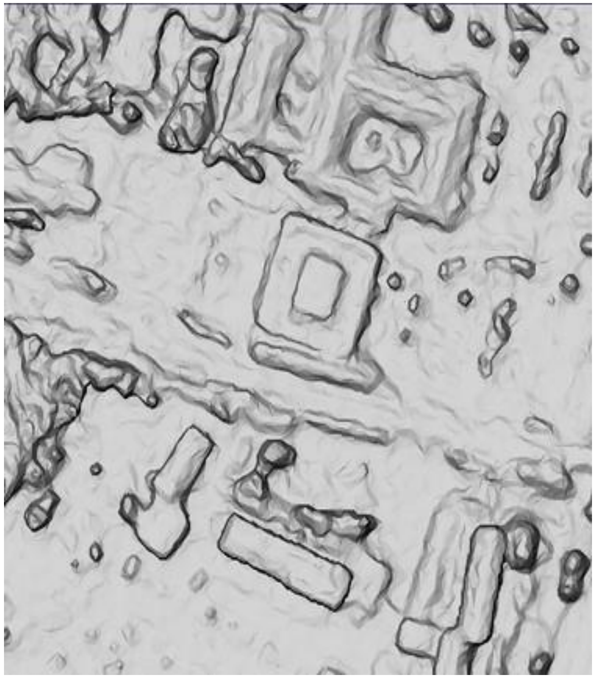
Raw Bands

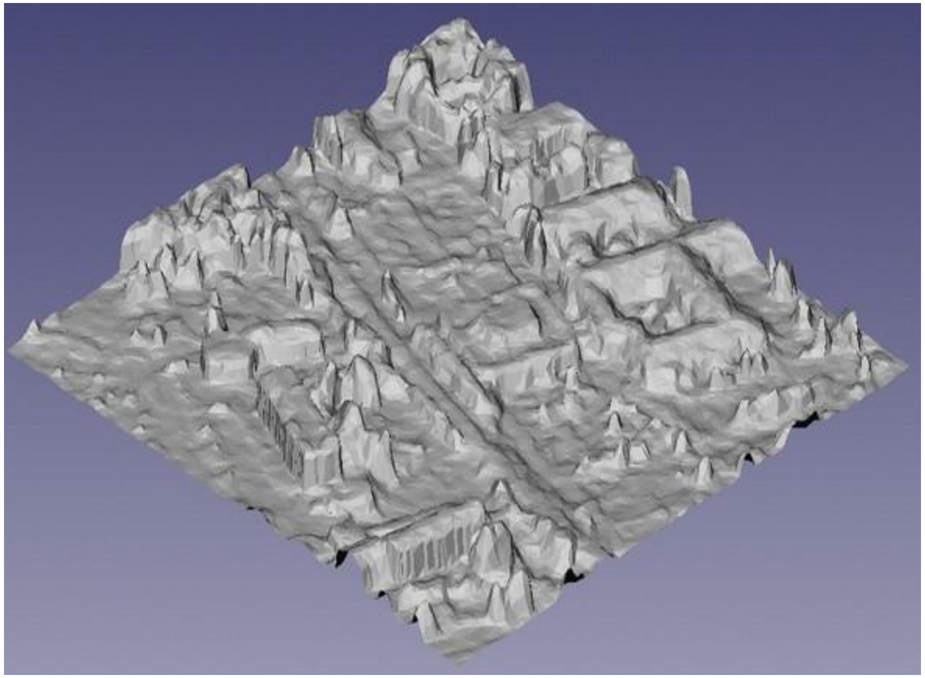
- Panchromatic [450-800] – blend of visible light into a grayscale;
- Coastal [400-450] – violet and deep blue, useful primarily for shallow waters, aerosols, dust, smoke;
- Blue [450-510] – useful for soil/vegetation discrimination, forest type mapping, and identifying man-made features;
- Green [510-580] – helps find oil on the surface of water, and vegetation (plant life); reflects more green light than any other visible color; man-made features are still visible;
- Yellow [585-625] – soils, sick foliage, hardwood, larch foliage (autumn);
- Red [630-690] - useful for identifying vegetation types, soils, and urban (city and town) features;
- Red Edge [705-745] – where the reflectance of vegetation changes rapidly;
- NIR1 [770-895] – used to measure plant health, good for mapping shorelines and biomass content; very good at detecting and analyzing vegetation;
- NIR2 [860-1040] – similar to NIR1
- + sharpened RGB



Collada

- Projected 3D mesh + textures.
- 13 levels of details, a lot of data, quadtree arrangement.



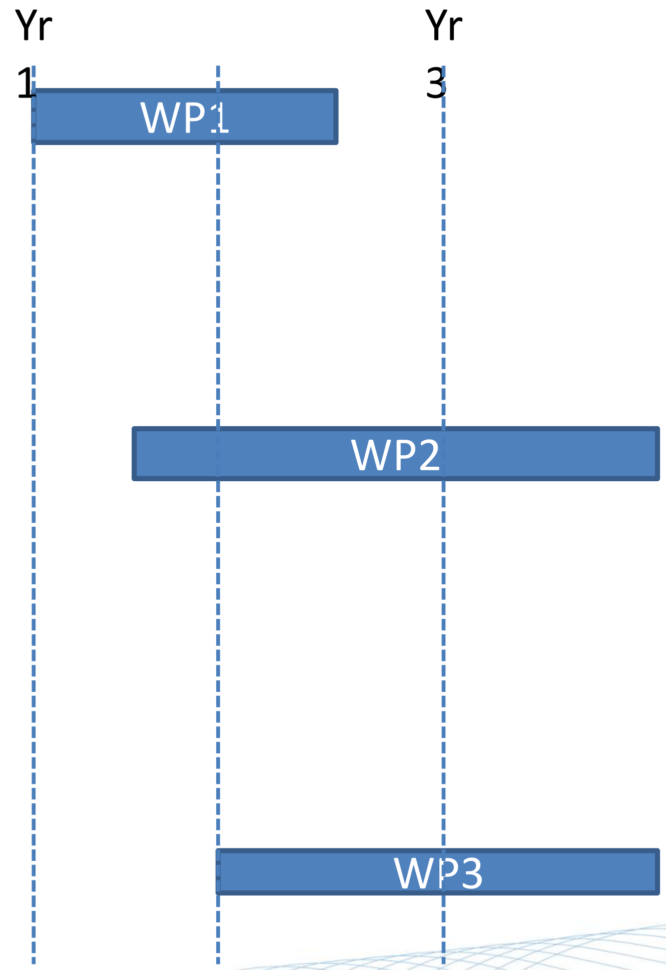


Project Workpackages

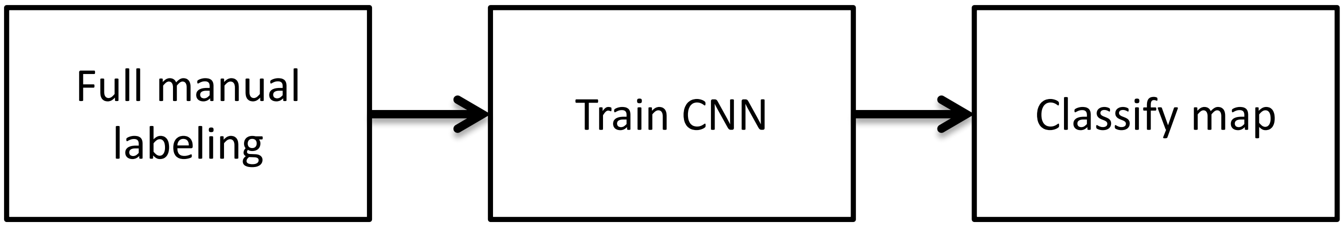
WP1 SEGMENT & CLASSIFY: segmentation and classification of 3D airborne map data

WP2 ANNOTATE & ANCHOR: segment regions in the map for classification and categorize the segments into fixed semantic categories.

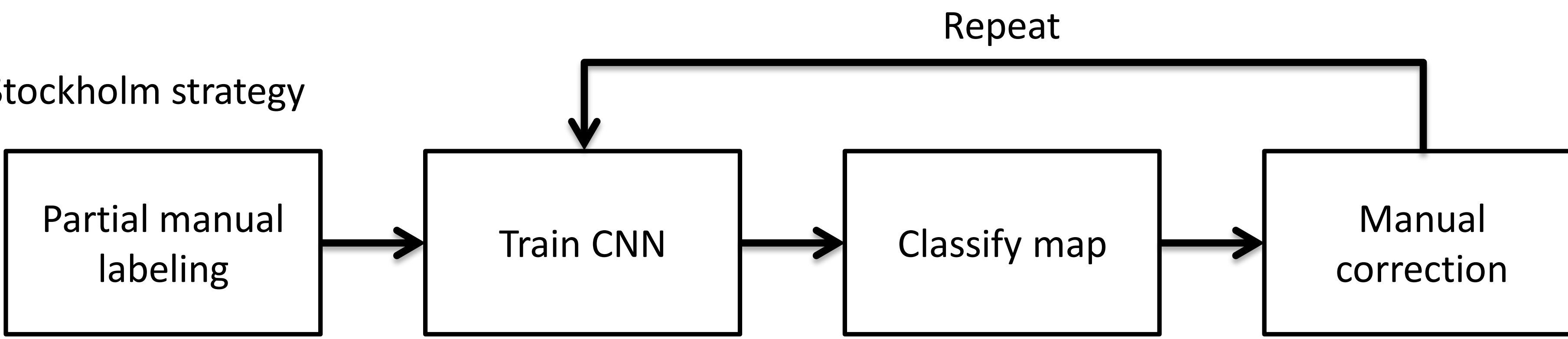
WP3 REASON & QUERY: a search engine which is composed of a reasoner and a query mechanisms



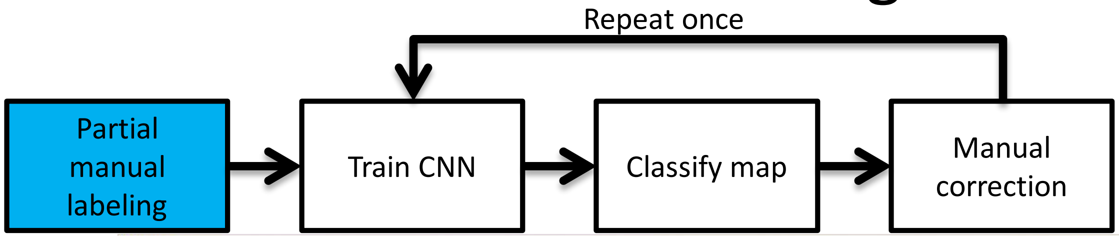
Traditional strategy

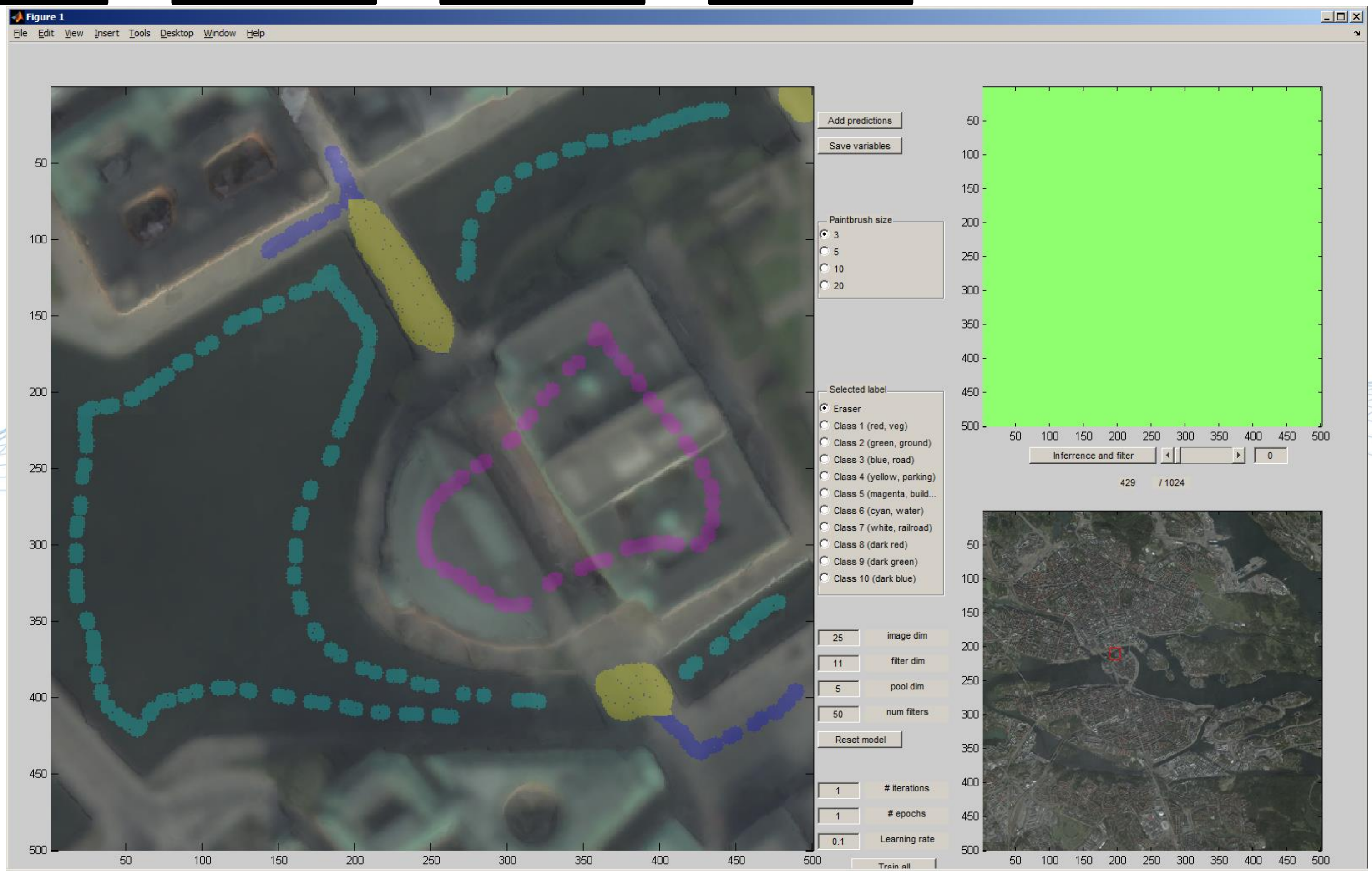


Stockholm strategy



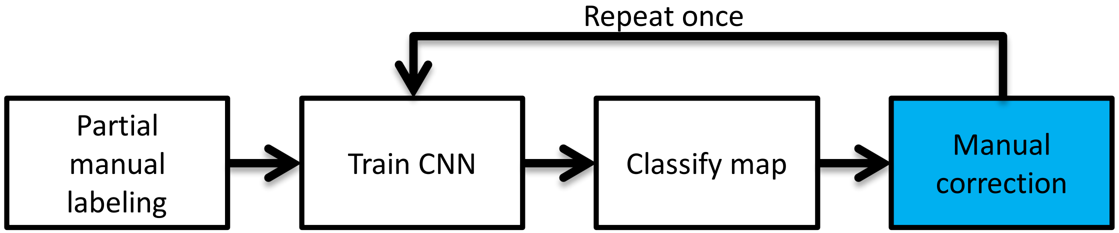
GUI for labeling and training

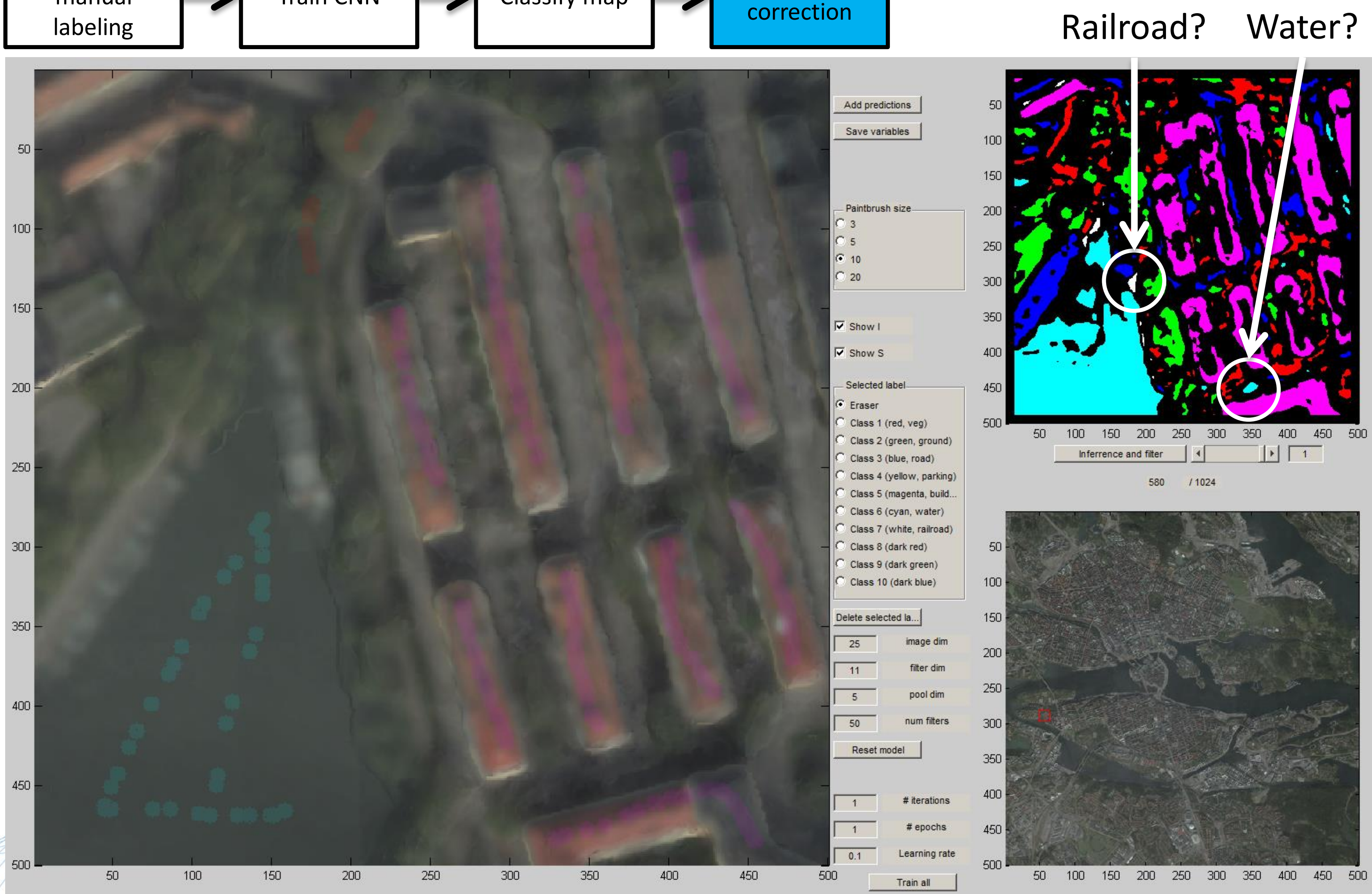




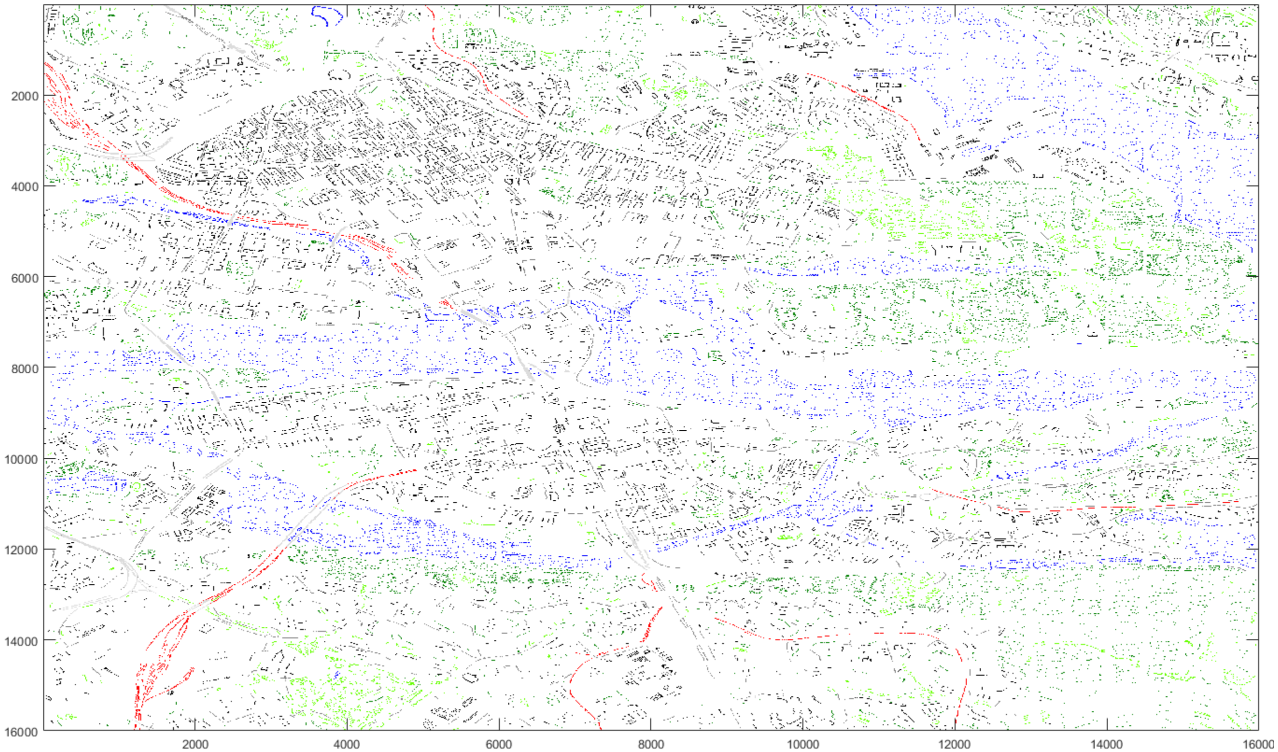
7 classes. *Parking* replaced by *bridge*.

Manual correction





Manually labeled groundtruth



Unlabeled	94.7%
Vegetation	1.1%
Ground	0.5%
Road	0.8%
Bridge	0.2%
Building	1.9%
Water	0.7%
Railroad	0.2%

5.3 % of map labeled in 6 hours. (100% would have taken >120 hours)

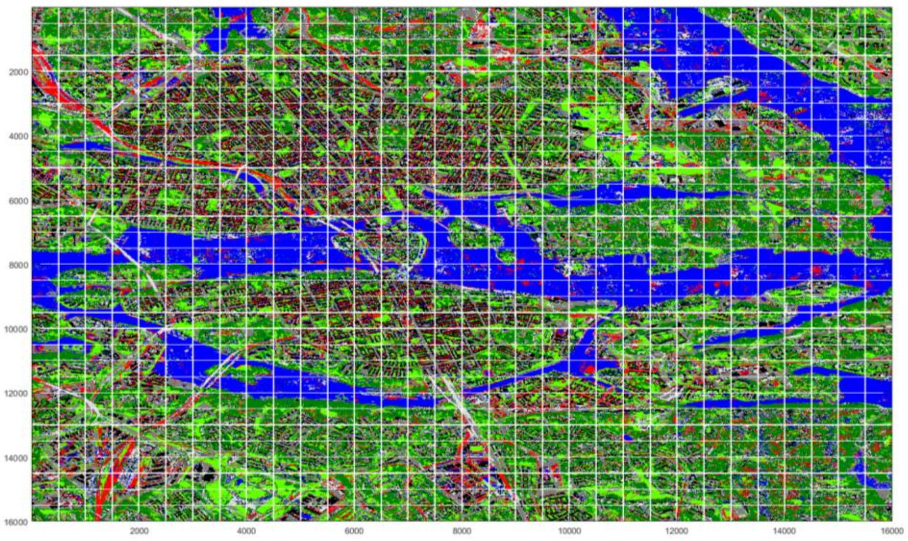
Classification result

- Trained on 90% initial manual labeled groundtruth
- Tested on 1000 randomly drawn pixels per class from the remaining 10 %

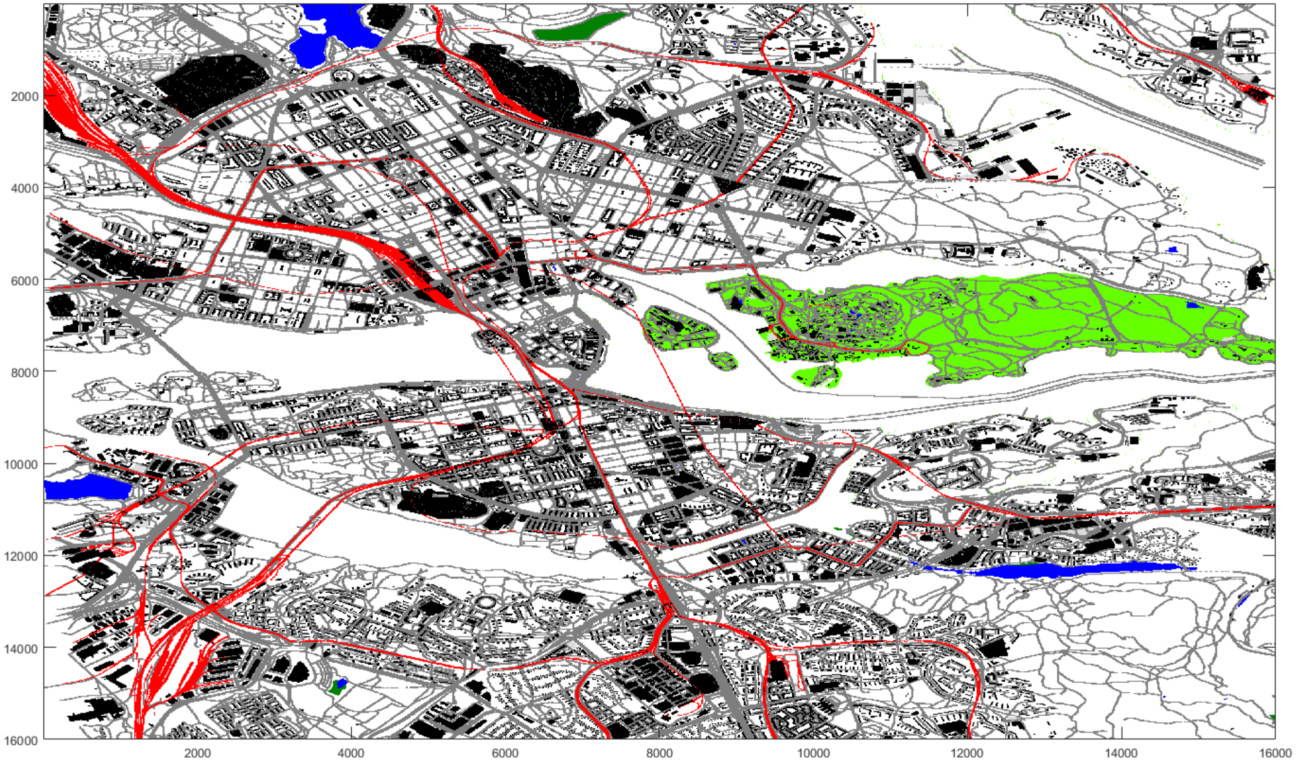
Equal class distribution in training data

95.6	1.1	0.3	1.4	0.3	0.9	2.9
1.2	**95.0**	0.6	0.7	0.1	0.8	1.7
0.2	1.3	**94.1**	3.4	0.1	0.4	3.8
0.7	0.7	1.9	**78.7**	7.2	2.4	3.6
0.2	0.1	0.1	9.2	**91.0**	0.4	0.9
0.5	0.7	0.5	2.7	0.6	**93.3**	2.4
1.5	1.1	2.5	3.8	0.6	1.9	**84.7**

Overall accuracy: 90.34%



OSM groundtruth



Unlabeled	52.4%
Vegetation	0.1%
Ground	2.8%
Road	28.6%
Parking	0.4%
Building	12.1%
Water	0.8%
Railroad	2.8%

Roads and buildings are well labeled, but not the other classes

Lantmäteriet groundtruth

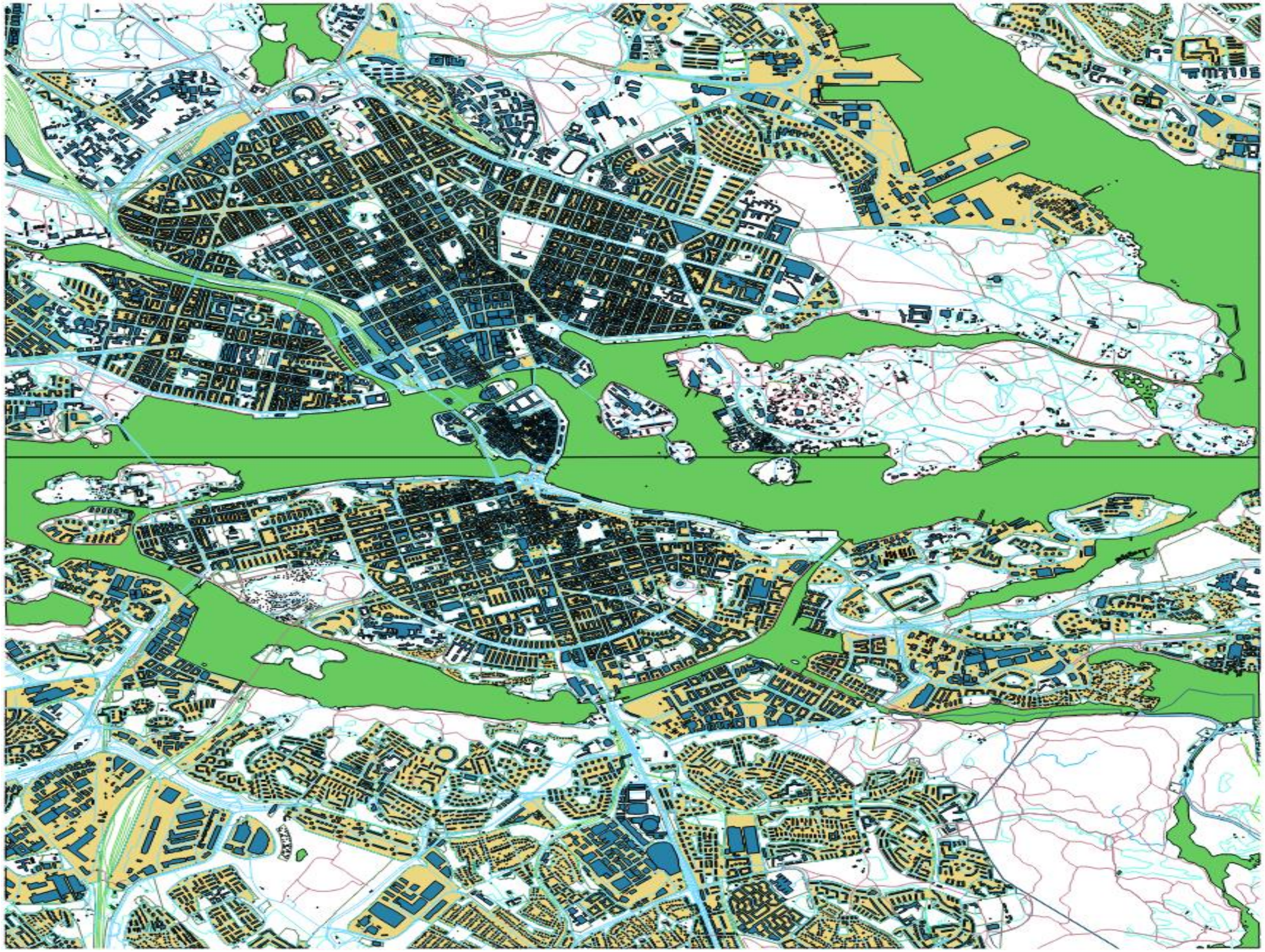
Unlabeled	17.8 %
Vegetation	11.4 %
Ground	15.4 %
Road	22.4 %
Bridge	0.0 %
Building	13.8 %
Water	17.0 %
Railroad	2.23 %

All classes well labeled, except bridge

Using **on-line** sources of geo-data

(I) Using **on-line** sources of geo-data:

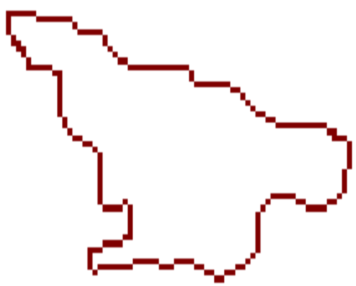
Läntmateriet (vector-based representation of geo-related features)

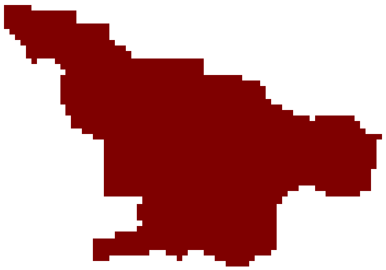


- Water (32)
- Road (8202)
- RailRoad (634)
- Building (20094)
- Industry
- Residential
- School
- Hospital
- Police Station
- Train Station
- City Hall
- Sport Hall
- Fire Station
- University
- ...

Läntmateriet (vector-based representation of geo-related features)

- Transforming the **vector-based** representation into a **raster-based** representation

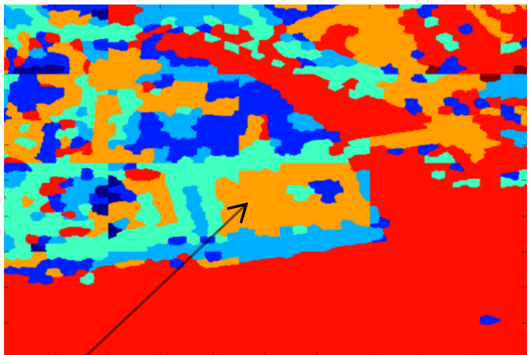




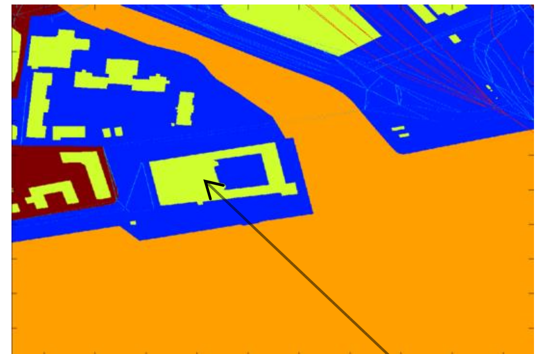
- *Published in **GIScience 2016** - Workshop on Spatial Data on the Web*

(I) Using **on-line** sources of geo-data:

Merging Läntmateriet with the Classifier output *(focusing on buildings)*



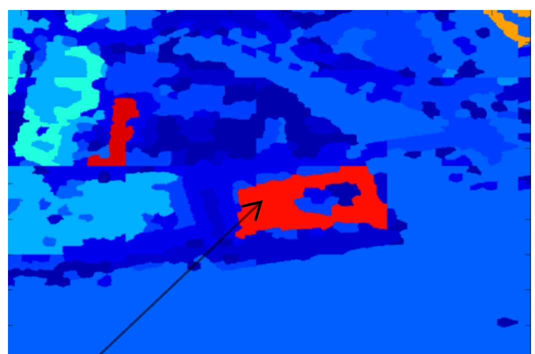
Classifier's output



Läntmateriet

Building

City Hall



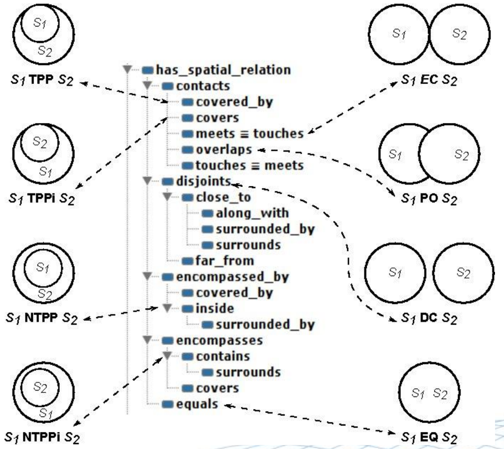
Enriched labels

City Hall

Semantic Modelling

Development of the domain knowledge of Boden in the form of an ontology (OntoCity)

- >100 000 objects stored in the ontology from the map
- Spatial Relations (along_with, close_to, …) RCC8
- Contains information about objects and their functions, affordances



Exploiting the **semantics**

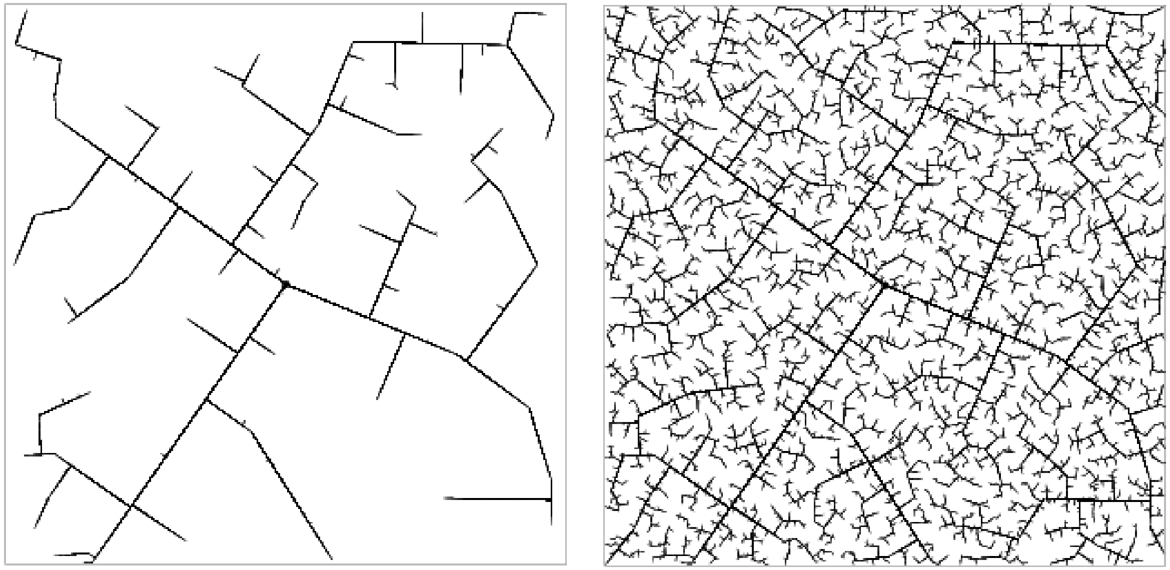
Aerial Path Planner Scenario:

- Scenario: simulating a drone flying over the city within a certain elevation range

- **RRT Path Planner** *(LaValle, Steven M. October 1998)*

  - RRT (Rapidly-Exploring Random Tree)

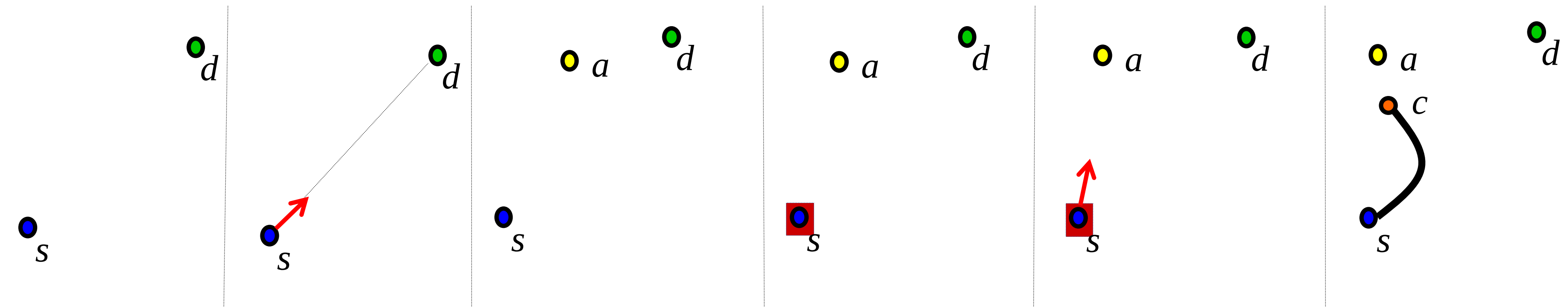
  - Is good at quickly finding a workable path



45 iterations 2345 iterations

- Scenario: simulating a drone flying over the city within a certain elevation range

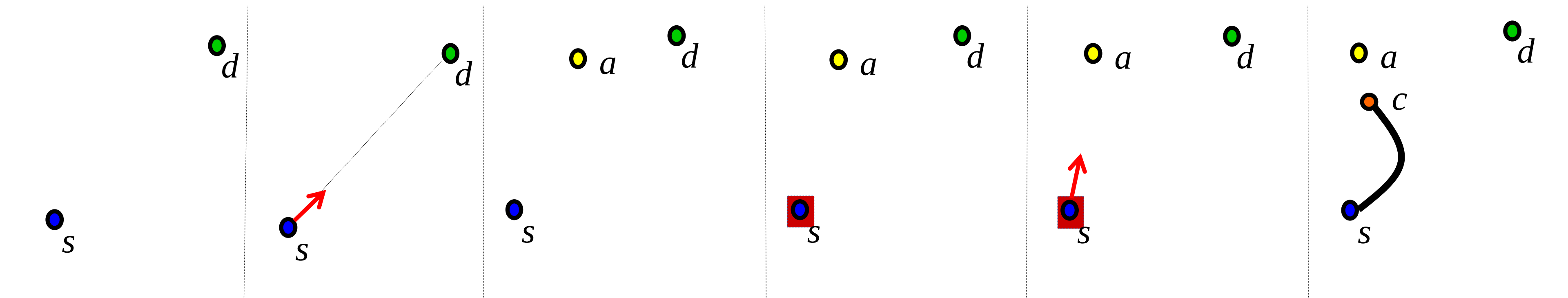
- **RRT Path Planner** *(LaValle, Steven M. October 1998)*



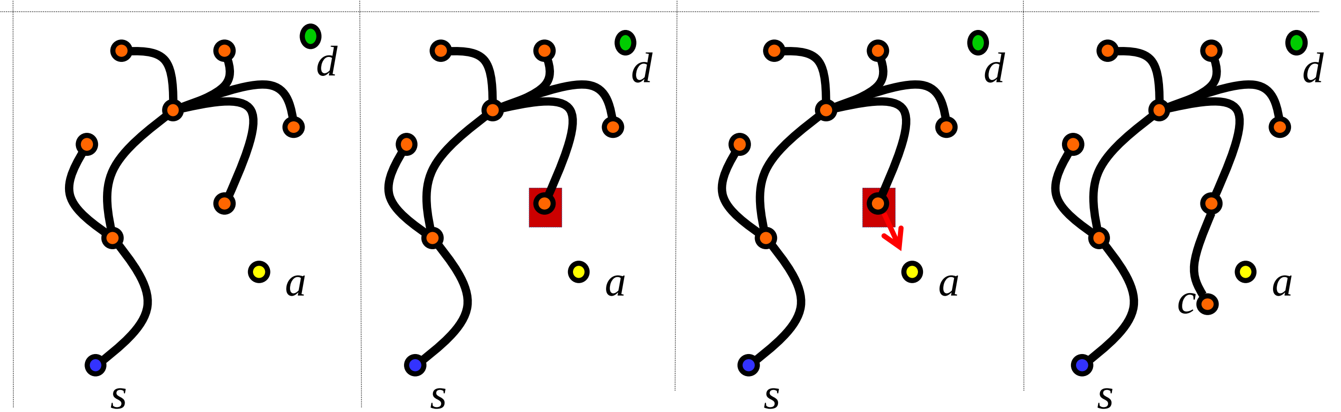
Add source to the tree	Check the direct connection to destination	Pick a **random** point *a* in space	Find a node in the tree which is closest to *a*	Drive towards a from the node in the tree	If path to *c* is **collision-f** add *c* to the tree

- Scenario: simulating a drone flying over the city within a certain elevation range

- **RRT Path Planner** *(LaValle, Steven M. October 1998)*



Add source to the tree	Check the direct connection to destination	Pick a **random** point *a* in space	Find a node in the tree which is closest to *a*	Drive towards a from the node in the tree	If path to *c* is **collision-f** add *c* to the tree



→**Collision Checking**

✓Based on the **Elevation** values of the regions (avoiding the obstacles)

✓Based on the **Semantics** of the regions (→ a preliminary version implemented)

No-constraint (only elevation)

Avoiding hospitals

→**Collision Checking**

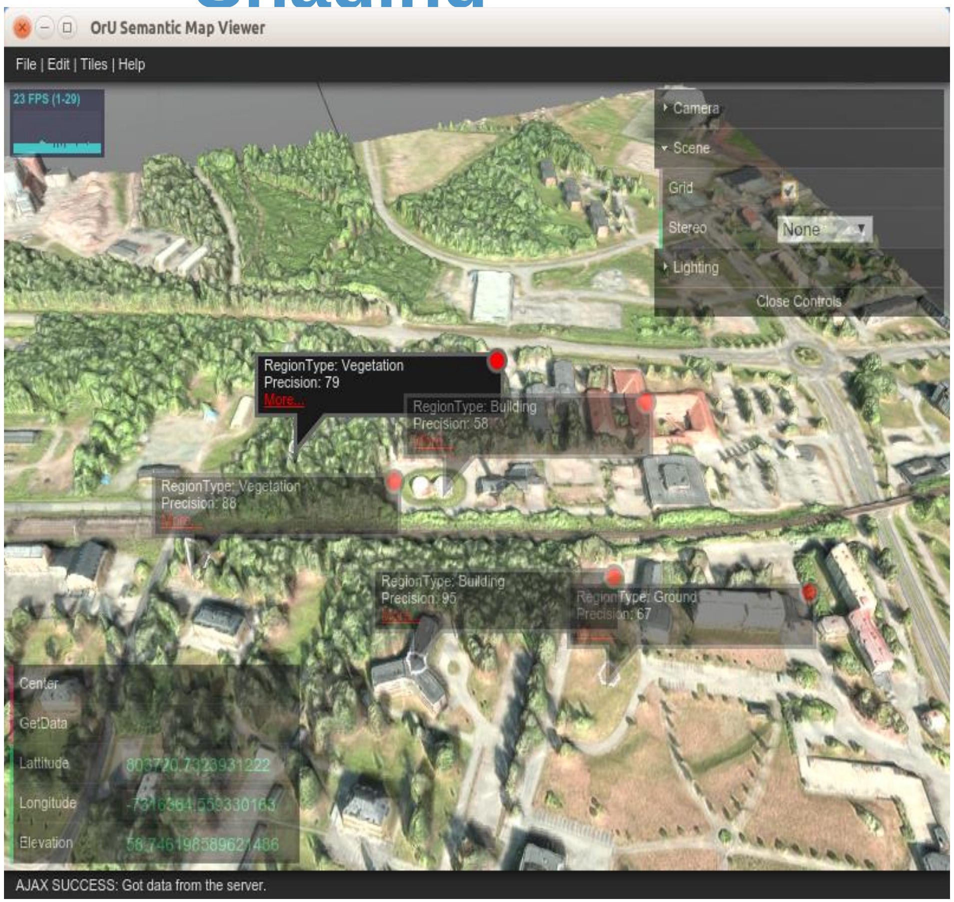
✓Since RRT is not complete, it may not find a path when there is semantic constraints

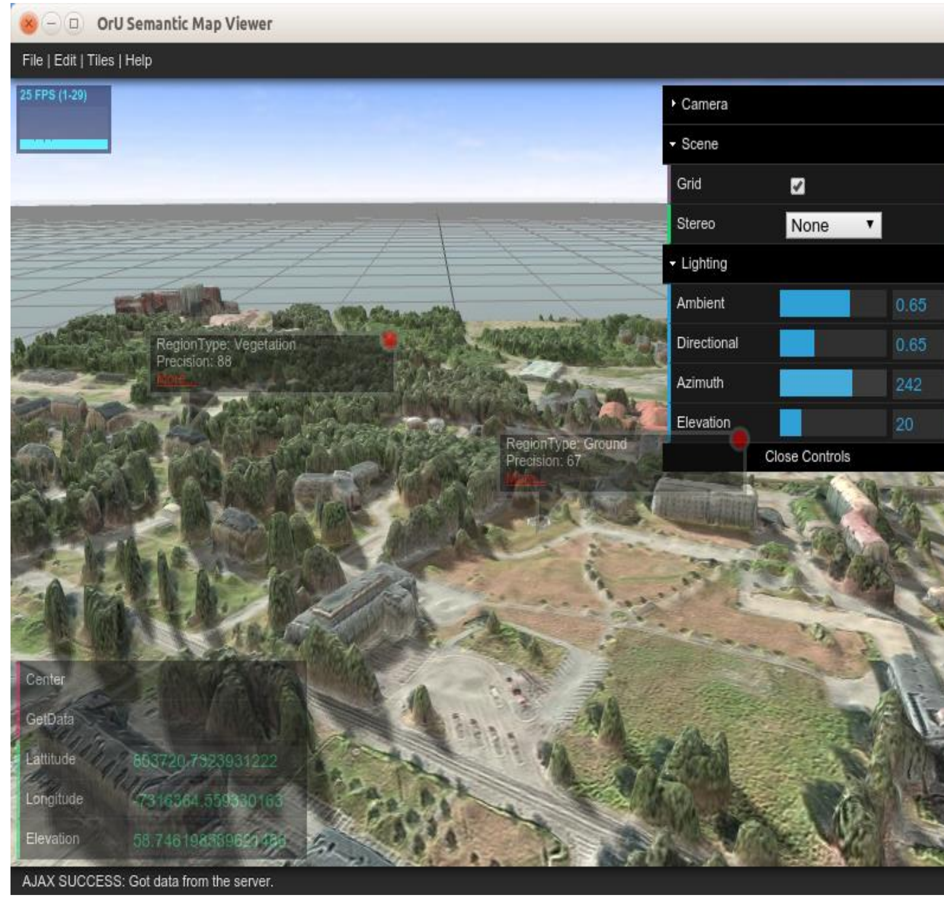
No-constraint (only elevation)

Avoiding Water

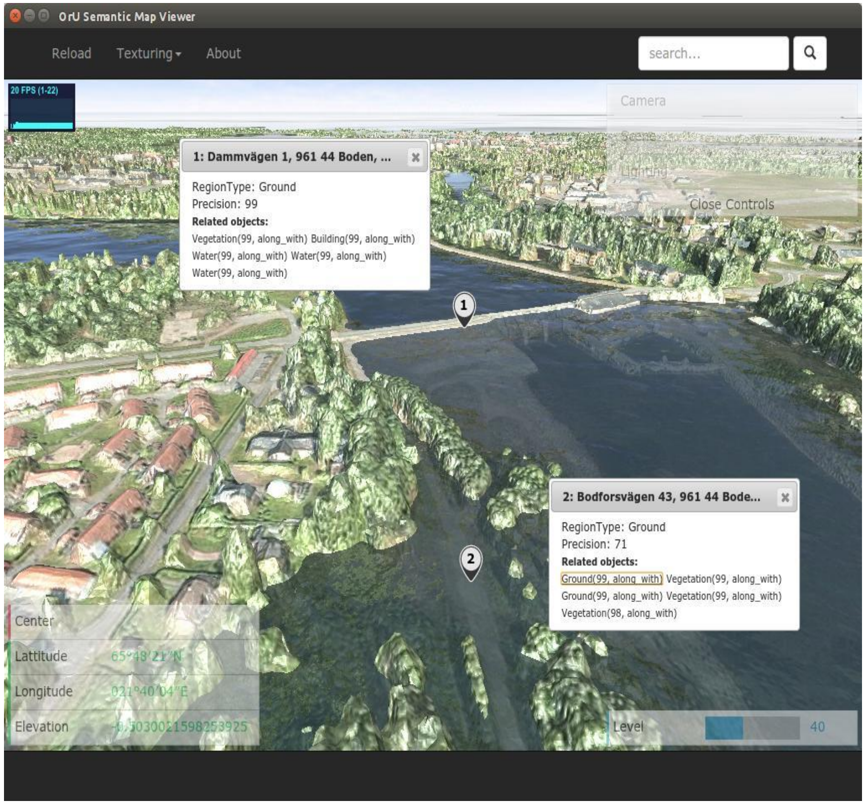
✓The **sampling process** needs to be guided by the **semantics**

SemMap Client 2.0: Labels and Shading



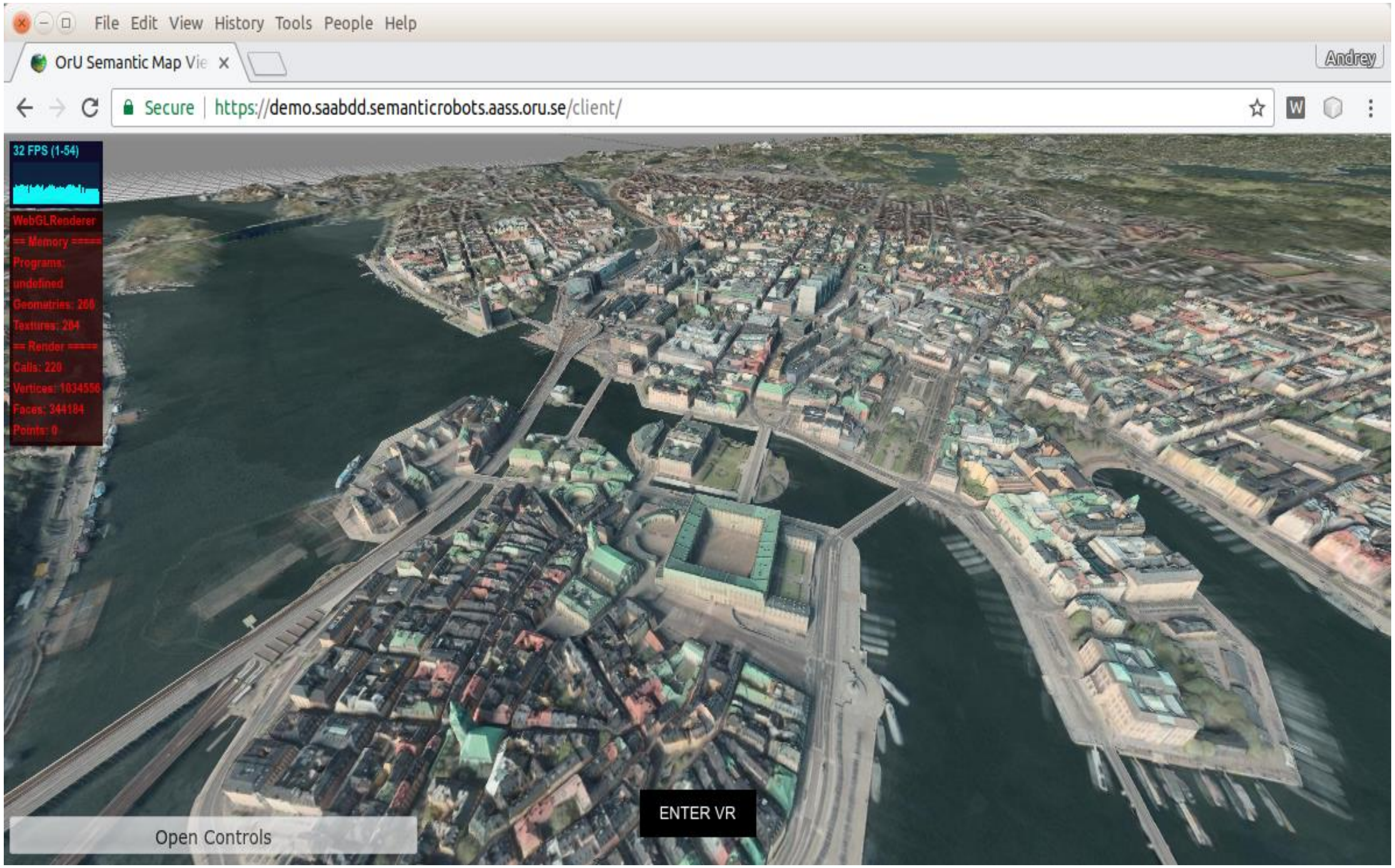


SemMap Client 2.1: Disaster Simulation

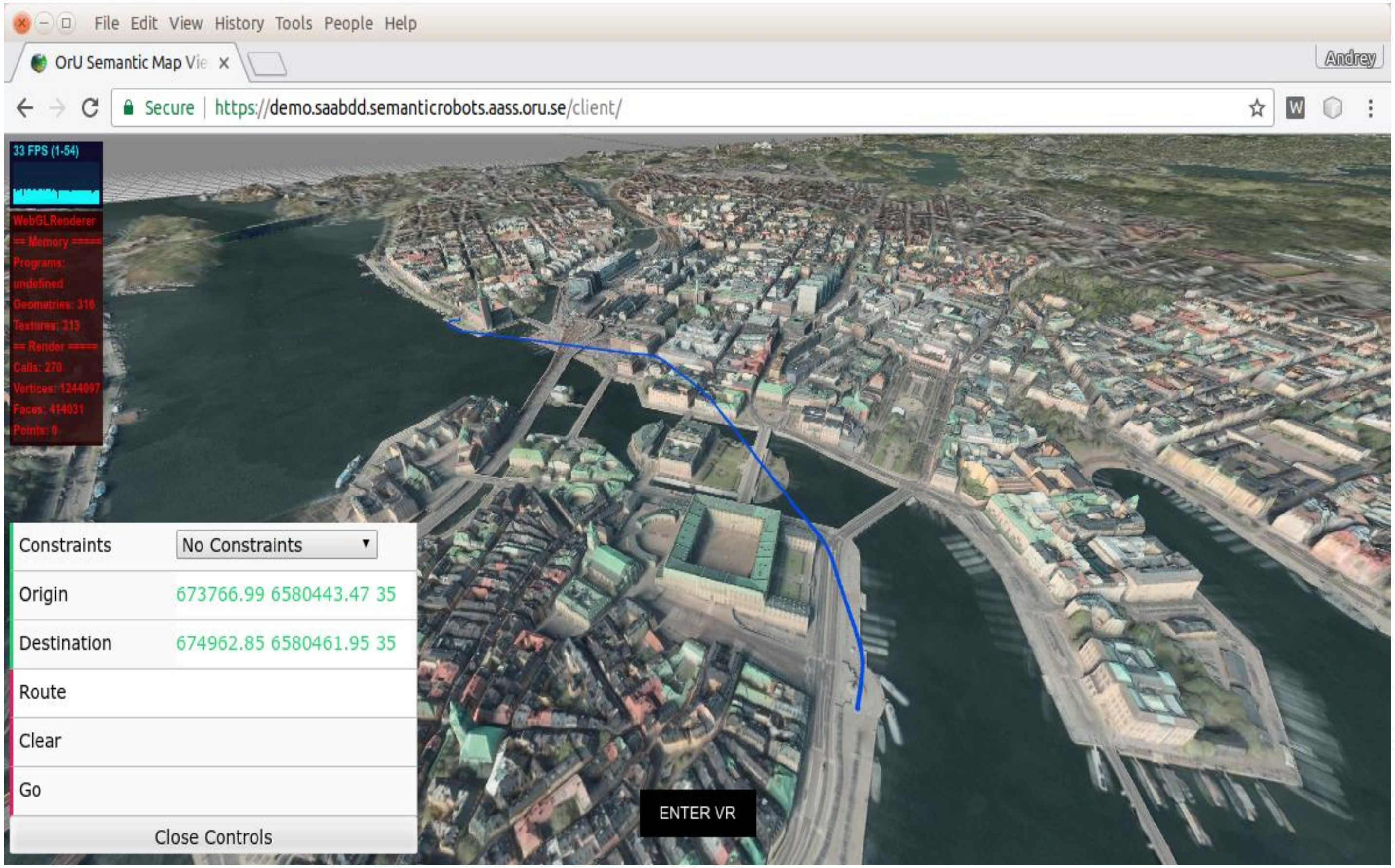


Flood simulation with the labels showing the PoI data when it is not flooded

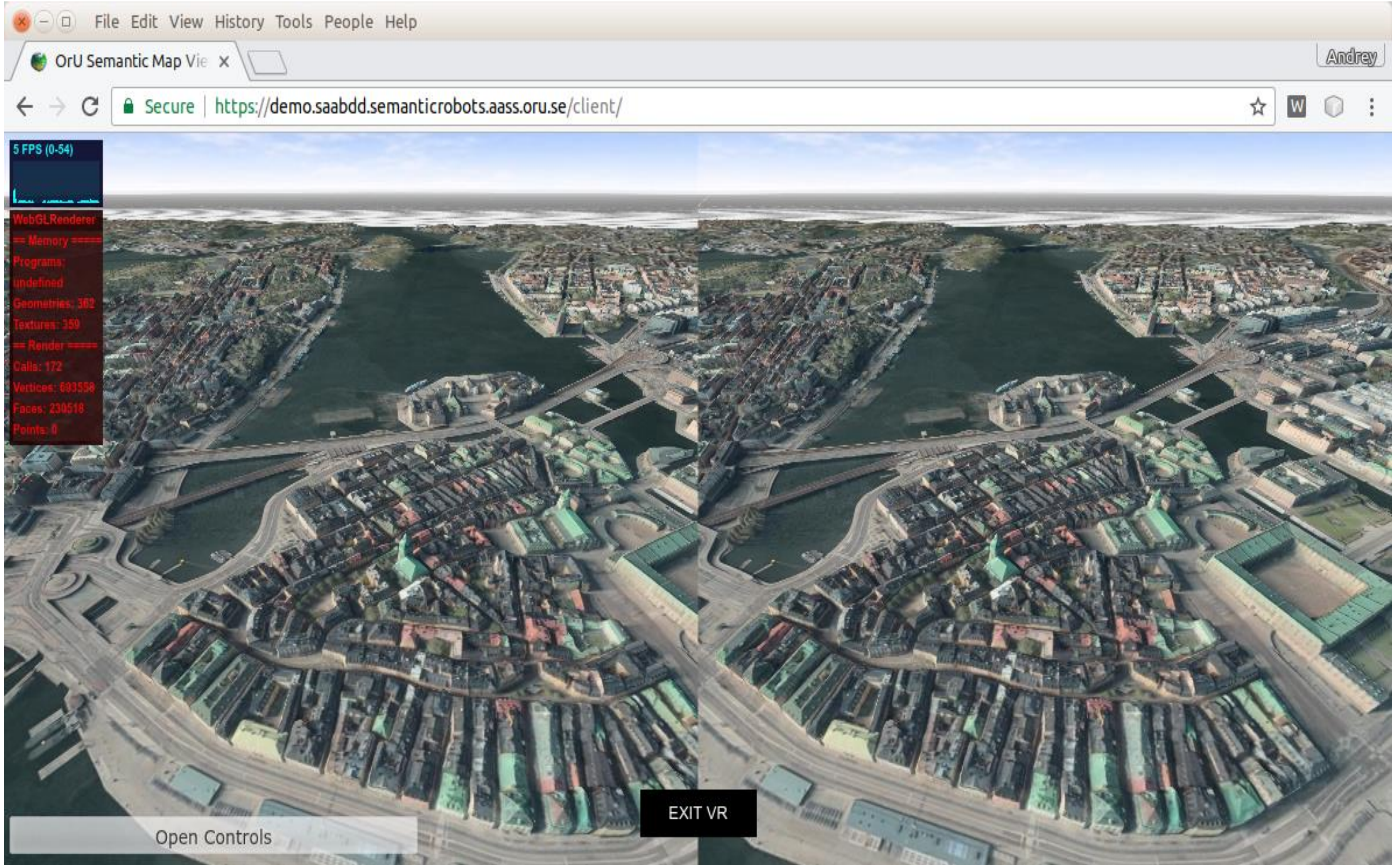
SemMap Client 3.0: Stockholm!



SemMap Client 3.0: Path Visualization

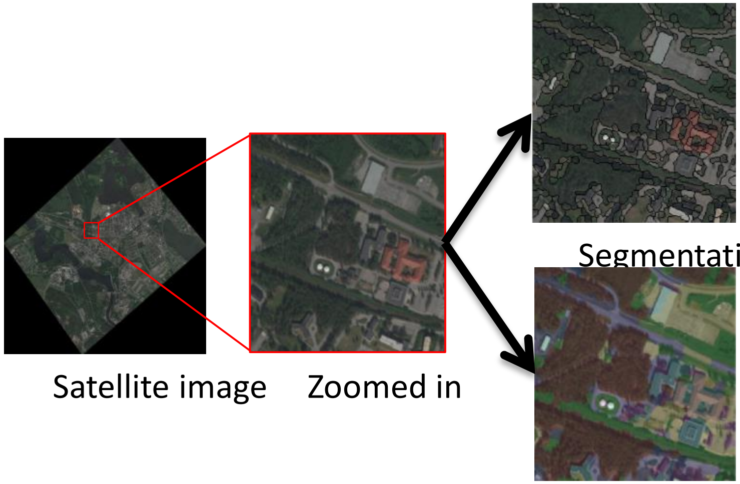


SemMap Client 3.0: VR



Full Pipeline from data to knowledge

WP1



Satellite image Zoomed in

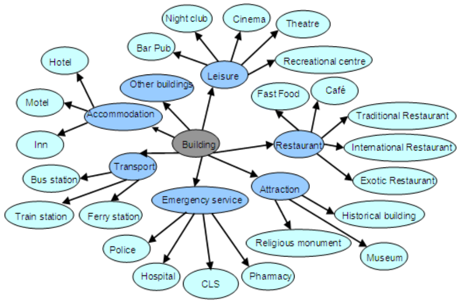
Segmentation

Classification

WP2

Object extraction and maintainance e.g. building 32, Address, functions, long/lat.

WP3



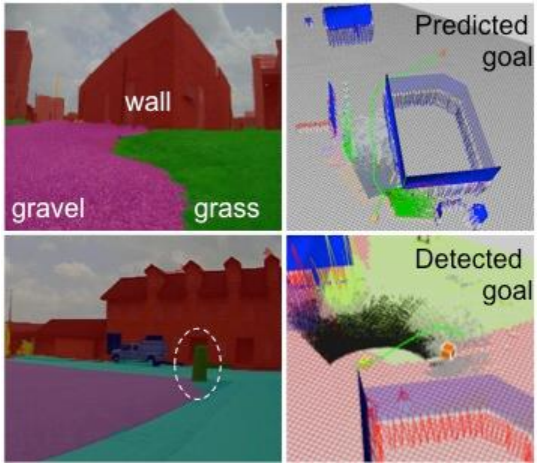
What is Semantic Perception

and why is it important

Process of augmenting sensor data into abstract, and typically, symbolic representations

It is important as it facilitates:

- Processes such as planning for robotics (e.g. Lift the crate)
- Integration of automated reasoning with sensor data



Semantic Perception requires meaningful semantics i.e. Shared and with relations between the concepts.

Thank you