SAND2015-2228C

Exceptional service in the national interest

Sandia National Laboratories

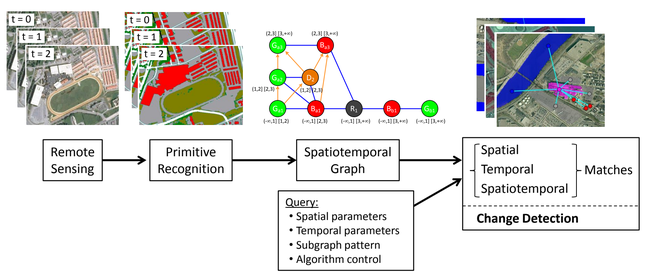
Image-Based Algorithms – Semantic Graph Algorithms

Randy Brost, Michelle Carroll, Will McLendon, Ojas Parekh, Danny Rintoul, David Strip, Diane Woodbridge
Sandia National Laboratories

Project goal: To make remote sensing data searchable, emphasizing facility and proliferation examples.

Background

Computation Flow

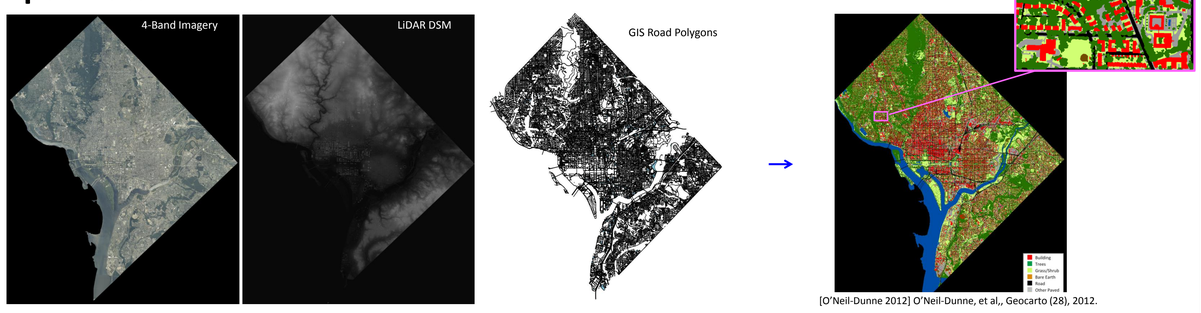


Graph representation:

- Analysis of objects and relationships.
- Is sensor agnostic.
- Allows multi-modality data.
- Encodes a variety of relationships: spatial, temporal, change, …
- Supports a variety of search algorithms.

Example Results (from MPD 2014)

Input Data



[O'Neil-Dunne 2012] O'Neil-Dunne, et al,, Geocarto (28), 2012.

Power Plant Search

Input:

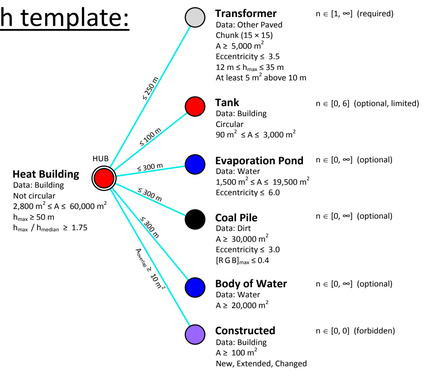
- Optical imagery
- LiDAR data
- GIS roads

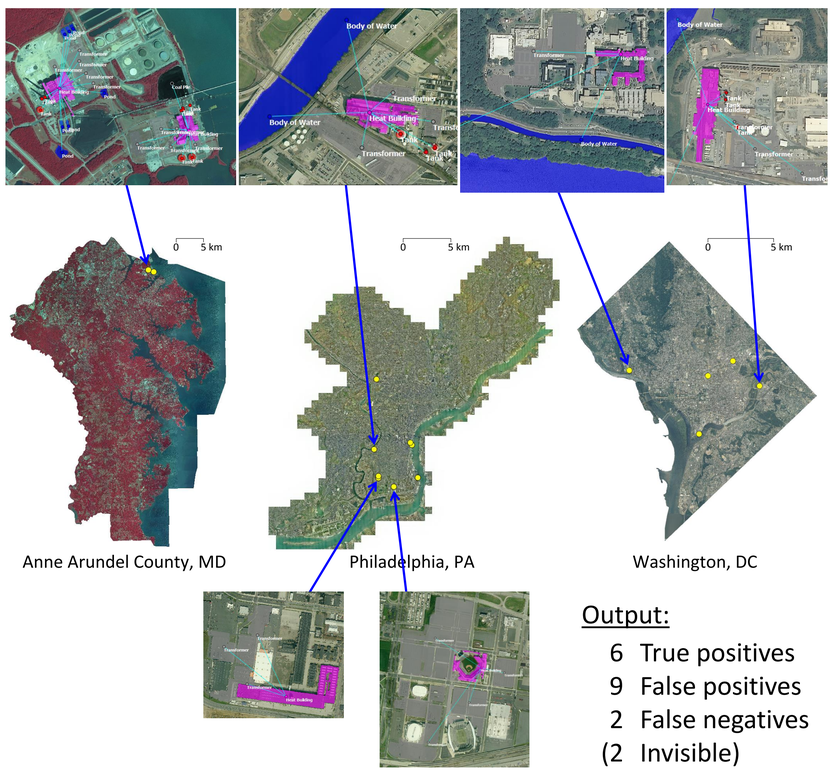
↓

- Land cover map

2,067 km² total area
135 billion Pixels
3.6 million Features

Search template:





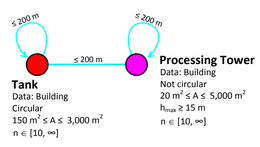
Anne Arundel County, MD — Philadelphia, PA — Washington, DC

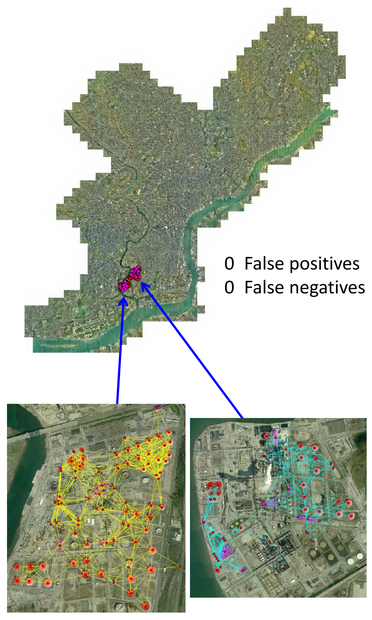
Output:

6 True positives
9 False positives
2 False negatives
(2 Invisible)

Large Refinery Search

Search template:





0 False positives
0 False negatives

Raw data points/pixels	101,495,378,523
Land cover pixels	8,981,933,044
Regions	1,133,822
Graph nodes	1,133,822
Buildings	154,062
Medium size buildings	87,170
Tank candidates	371
Tank complexes	28
Large refineries	2

Questions

- Robustness and scope of image pre-processing?
- Can we make it easier to construct queries?
- Search accuracy, in terms of both false positives and false negatives?
- Match quality score, ranking?
- Uncertainty characterization?
- Specific proliferation examples?

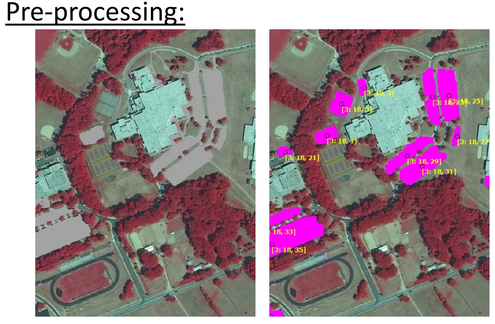
Advances

Evaluation/Publication

- Measurement of False Negatives by expert image analyst using government databases.
- Brost, *et al*, "A computational framework for ontologically storing and analyzing very large overhead image sets," 3rd ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data (BigSpatial-2014), November 2014 (Best Paper Award at workshop).

Multi-Step Search

Pre-processing:

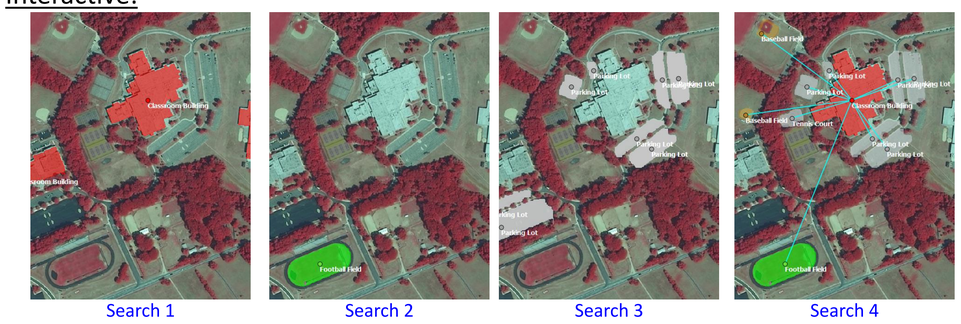


Search results can be written back to the graph, used as components of later searches.

Advantages:

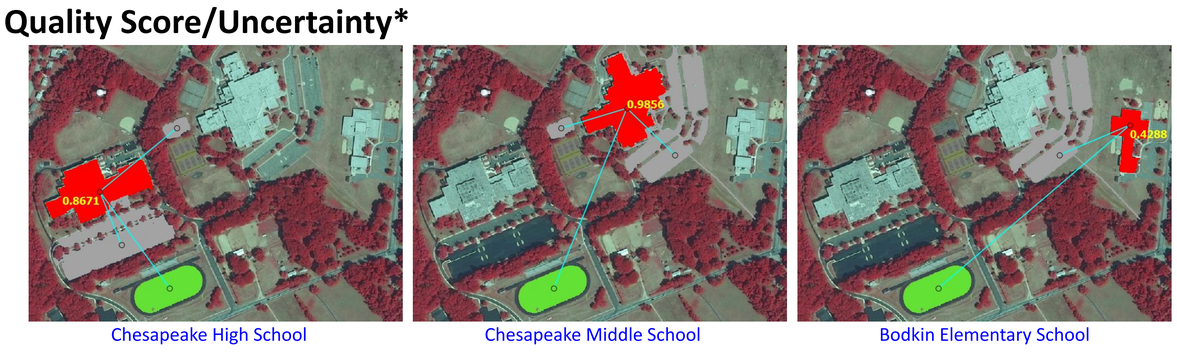
- Smaller, modular queries.
- Search re-use.
- Hierarchical semantics.
- Faster.

Interactive:

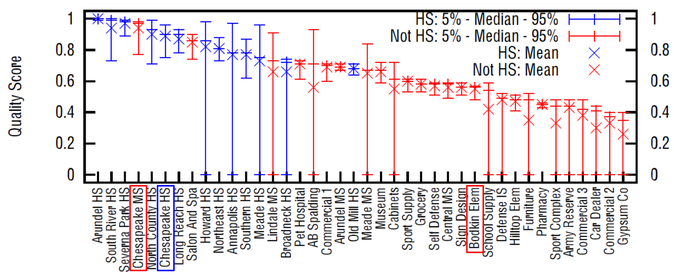


Search 1 — Search 2 — Search 3 — Search 4 (Using results from 1, 2, 3)

Quality Score/Uncertainty*



Chesapeake High School — Chesapeake Middle School — Bodkin Elementary School



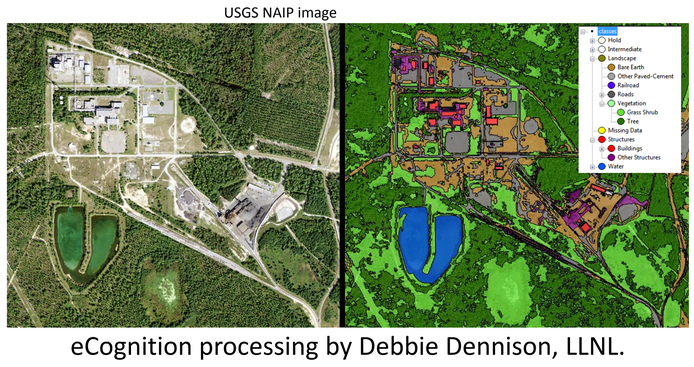
From Stracuzzi, Brost, Phillips, Robinson, Wilson, and Woodbridge, "Computing Quality Scores and Uncertainty for Approximate Pattern Matching in Geospatial Semantic Graphs," in review.*

* The theoretical foundations of this work, calculation of the plot, and submitted paper were supported by a related project supported with Sandia internal funding. NA-22 funding supports implementation of this method in the automatic search code, shown in the images and still in progress.

Image Pre-Processing

- UVM eCognition workshop SNL, including LLNL, LANL.
- eCognition, QT Modeler, Global Mapper.
- Obtained land cover data and rule sets from UVM, reproduced results.
- Pursuing general-purpose land cover rule sets through UVM analysis of Benchmark Imagery, internal efforts.
- See initial LLNL land cover results presented in Randy Roberts' talk.

USGS NAIP image



eCognition processing by Debbie Dennison, LLNL.

Application to Proliferation Examples

- Survey of prior proliferation sites (LANL).
- Generation of ontology and GeoQuestion search template based on ontology (LANL).
- Collection of site imagery (LLNL).
- Generate graph based on processing site imagery (SNL).
- Execute search based on ontology (SNL).
- Thorough multi-site search, refinement, and evaluation planned (All).

U.S. DEPARTMENT OF ENERGY
National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND No. 2011-XXXXP

Image data provided by the University of Vermont unless otherwise noted.