VISUAL SEARCHES IN VERY LARGE DATA SETS USING DEEP LEARNING TECHNIQUES

AU GIS DAY

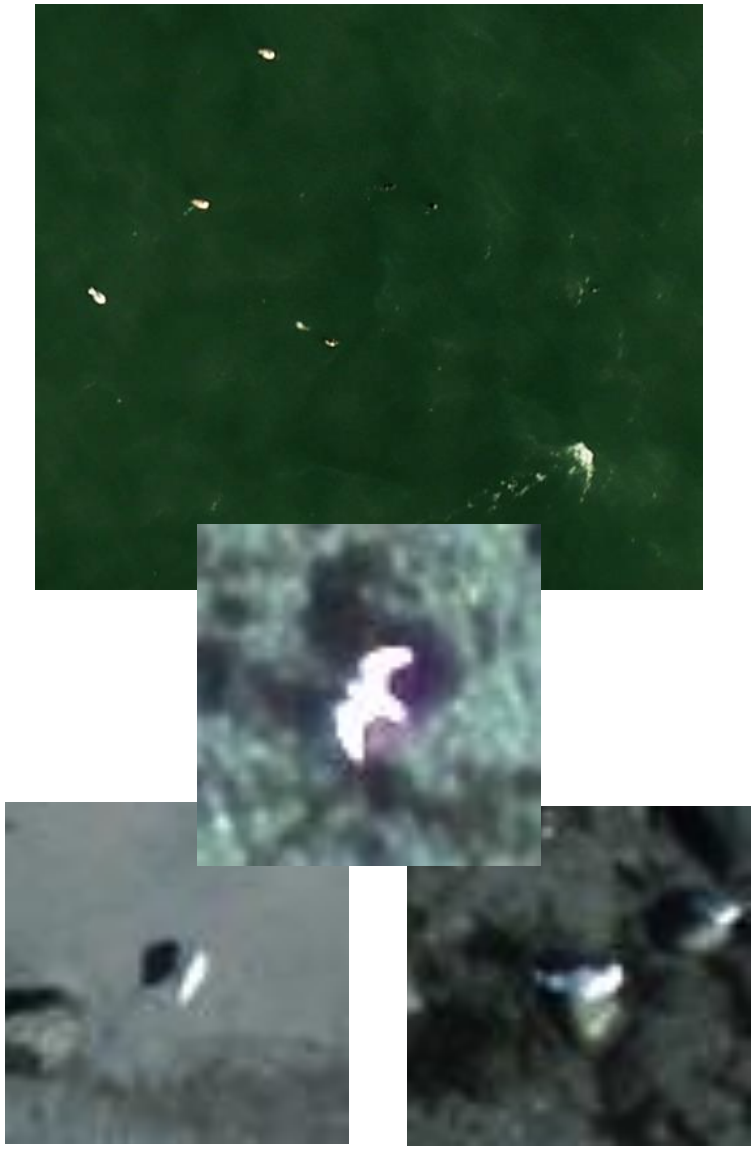
Karsten Østergaard Noe

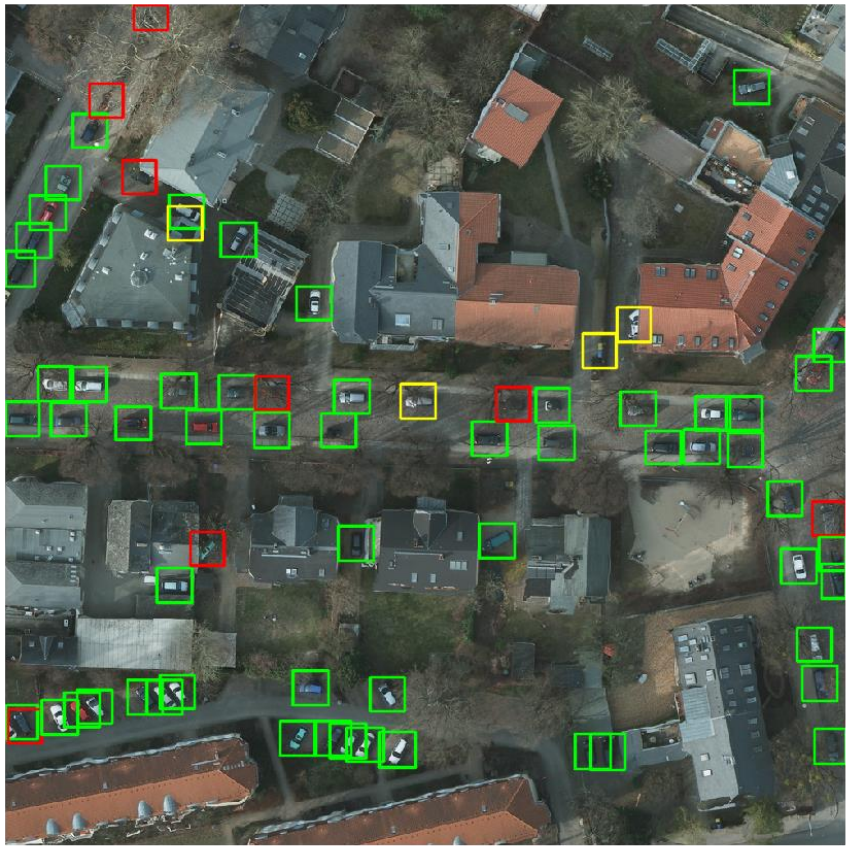
Principal Visual Computing Specialist, The Alexandra Institute

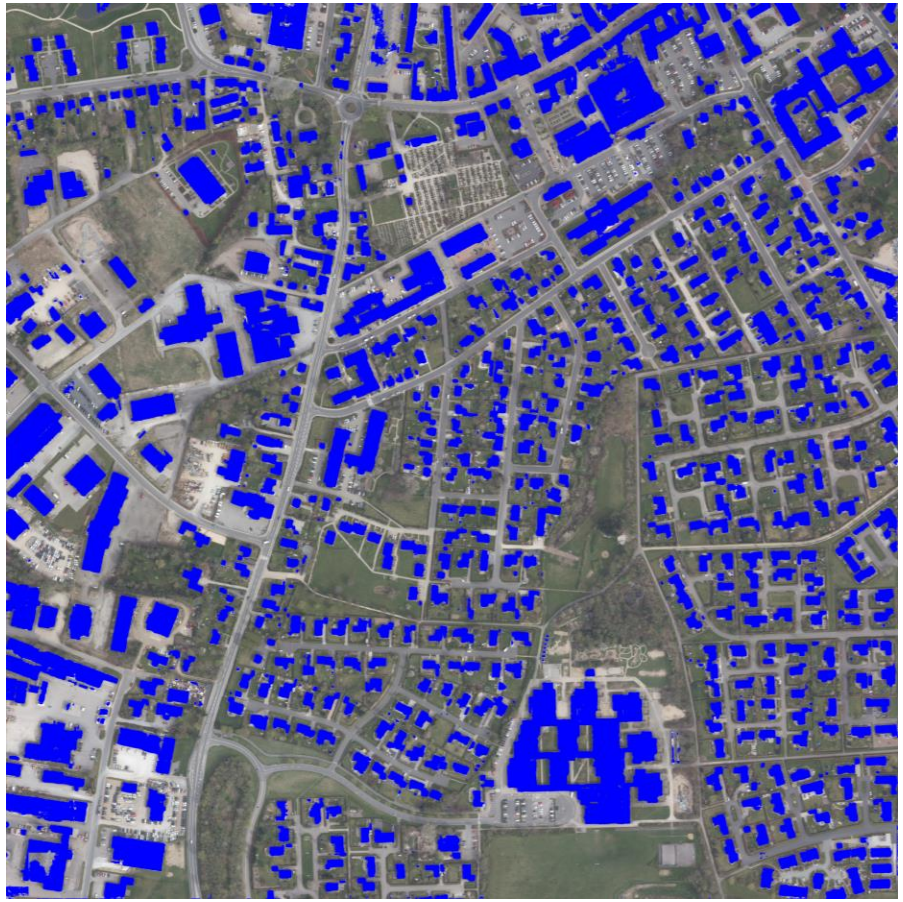
karsten.noe@alexandra.dk

ALEXANDRA INSTITUTTET

DEEP LEARNING FOR CLASSIFICATION, OBJECT DETECTION AND SEGMENTATION



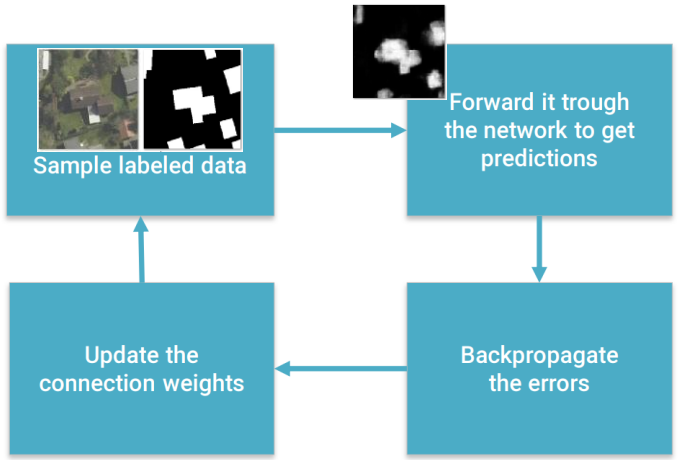




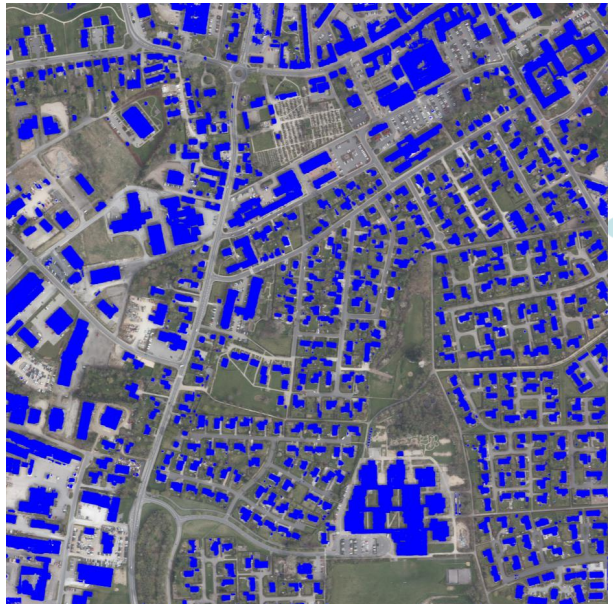
ALEXANDRA INSTITUTTET

Deep learning

- Supervised learning
  - Need for annotated training data
  - Time consuming / costly
- Deep learning only works when we have a lot of training examples
- Can we develop methods that don't require training data?



FIELDSENSE



Agency for Data Supply and Efficiency

ALEXANDRA INSTITUTTET

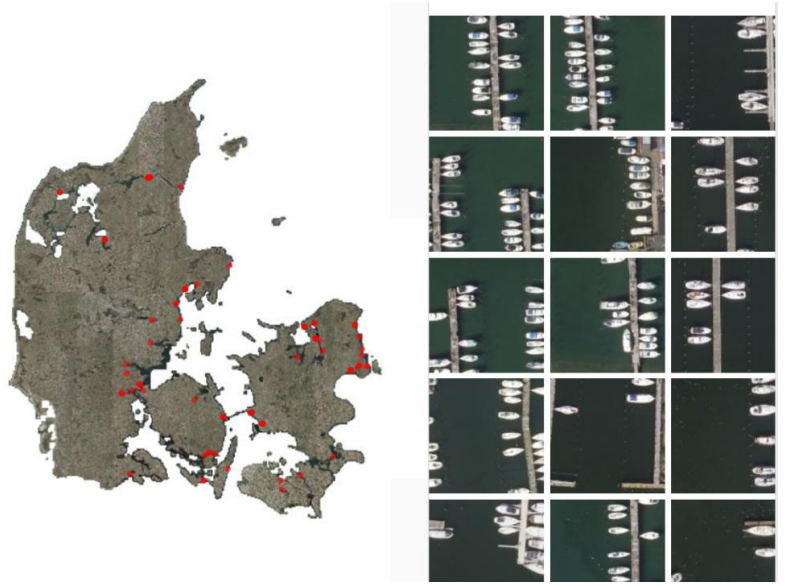
GEOVISUAL SEARCH

- Inspired by Descartes Labs' geovisual search (http://search.descarteslabs.com)
- *So ein Ding müssen wir auch haben!*
- We developed our own version.
- Working with Danish data sets
  - Free orthophotos (kortforsyningen.dk)

ALEXANDRA INSTITUTTET

POTENTIAL USES FOR GEOVISUAL SEARCH

- Finding objects in large areas: *Where can I find marinas?*
- Segmentation of large areas
  - Land use, habitat types
- Change detection
  - Comparing descriptors with different time stamps

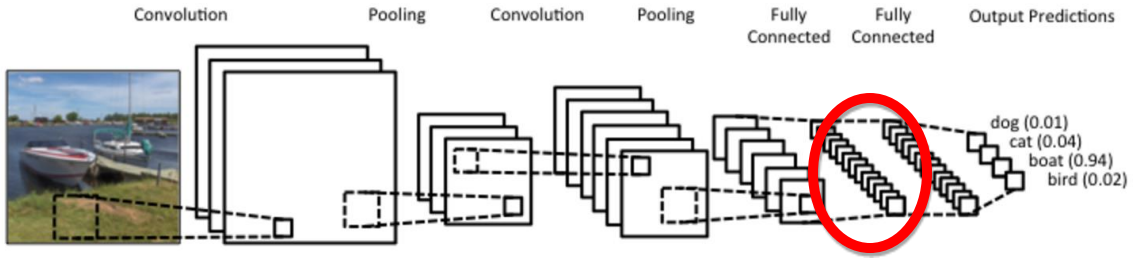


- **Fast annotation of training data for machine learning**

ALEXANDRA INSTITUTTET

GEOVISUAL SEARCH: THE TECHNOLOGY

- Orthophotos over Danmark divided into 48 million cutouts of 224*224 pixels (20 cm/pixel)
- A resnet-34 neural network used to extract 2048 numbers per cutout (393 GB)
  - We call this a descriptor

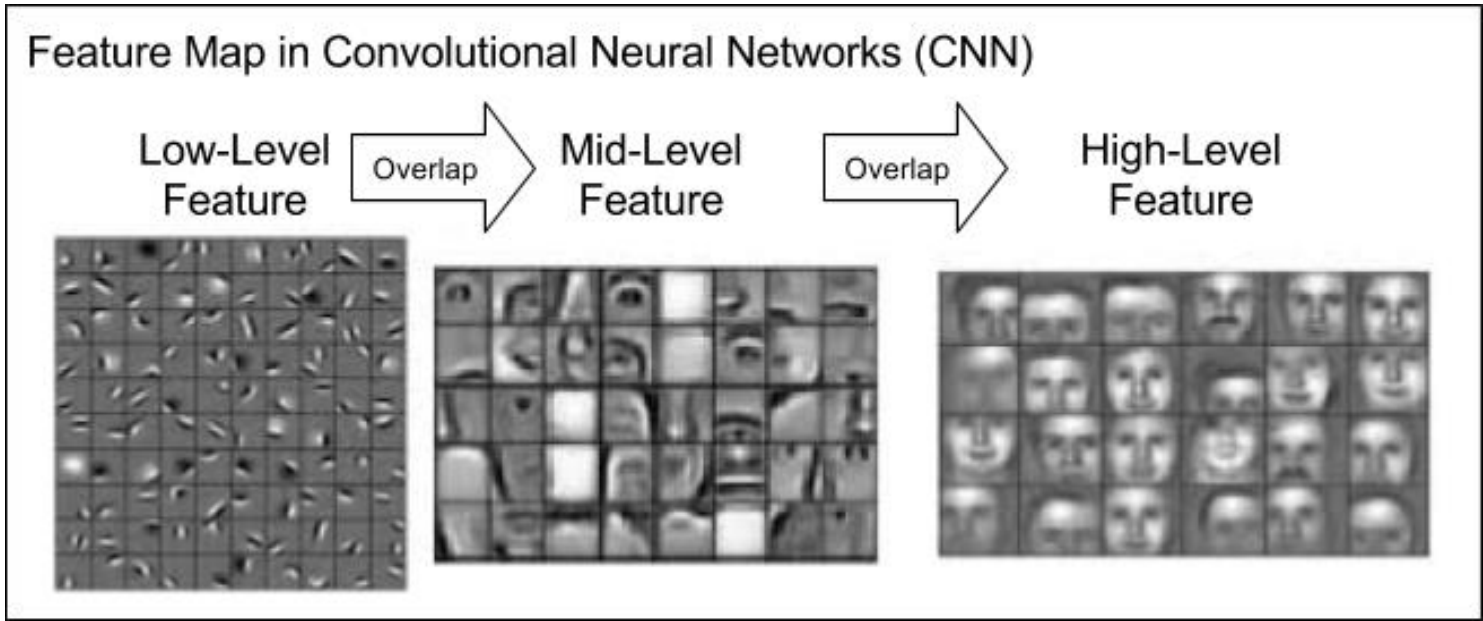


  - Another network (autoencoder) reduces each descriptor to 512 bits. (3 GB)
- When an area is clicked the most similar among the 48 mio descriptors are found and returned. This search takes less than 80 ms.

ALEXANDRA INSTITUTTET

USING A CNN TRAINED ON NATURAL IMAGES

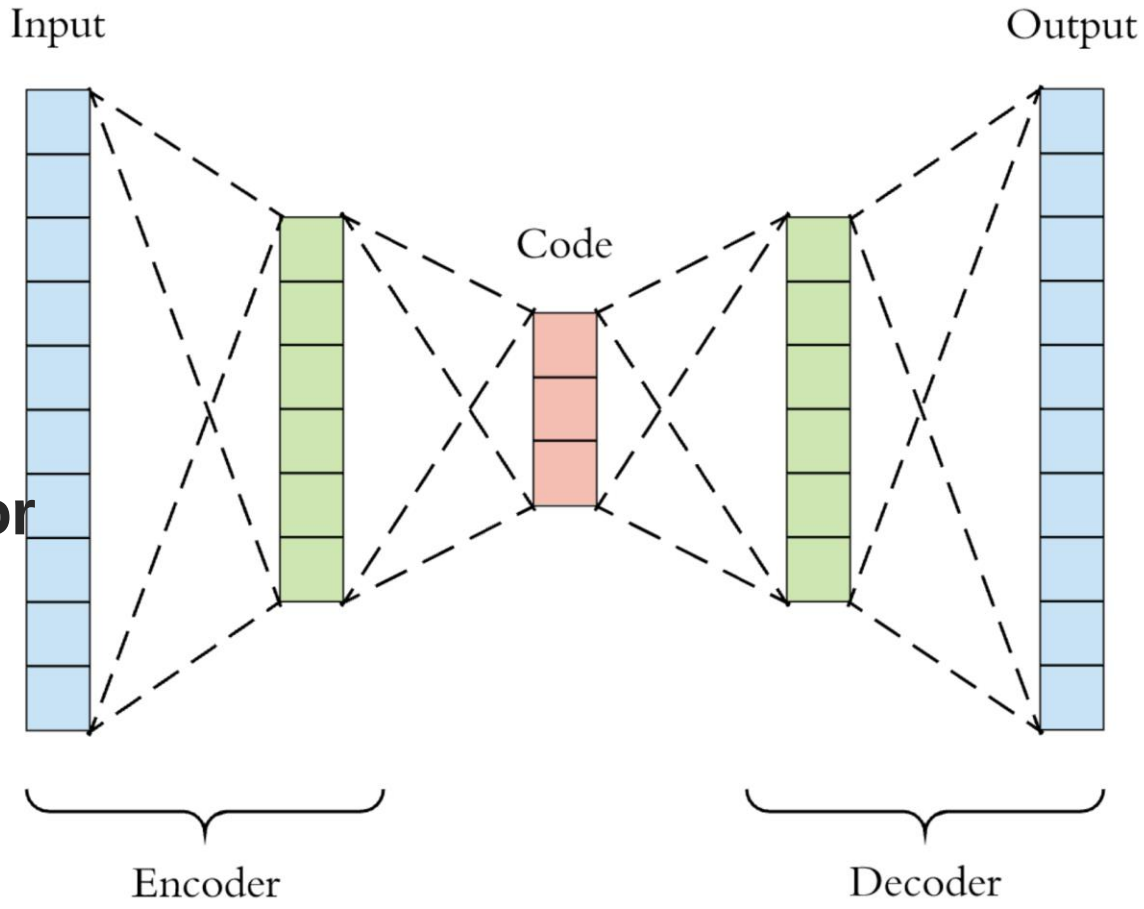
- The used network is pretrained on ImageNet (1000 classes)
  - Natural images (cats, dogs, cars etc.)
- The 2048 numbers describe responses to trained filters.
- Represents a way of comparing images: The distance between two descriptors correlates to the probability of the two images containing the same object.



AUTOENCODER

- Compression technology
- Neural network encodes 2048 values into 512 values
- *Binarization constraint*
- **Extracts components from descriptor relevant for the orthophoto**

- Now the comparison is done using Hamming distance



ALEXANDRA INSTITUTTET

GEOVISUAL SEARCH: EKSEMPLER

Try it here: ***denmark3d.alexandra.dk/geosearch***

IMPROVEMENTS NOT IN THE WEB VERSION

- Different cutout sizes / scales
- Better rotation invariance
- Interactive search refinement

ALEXANDRA INSTITUTTET

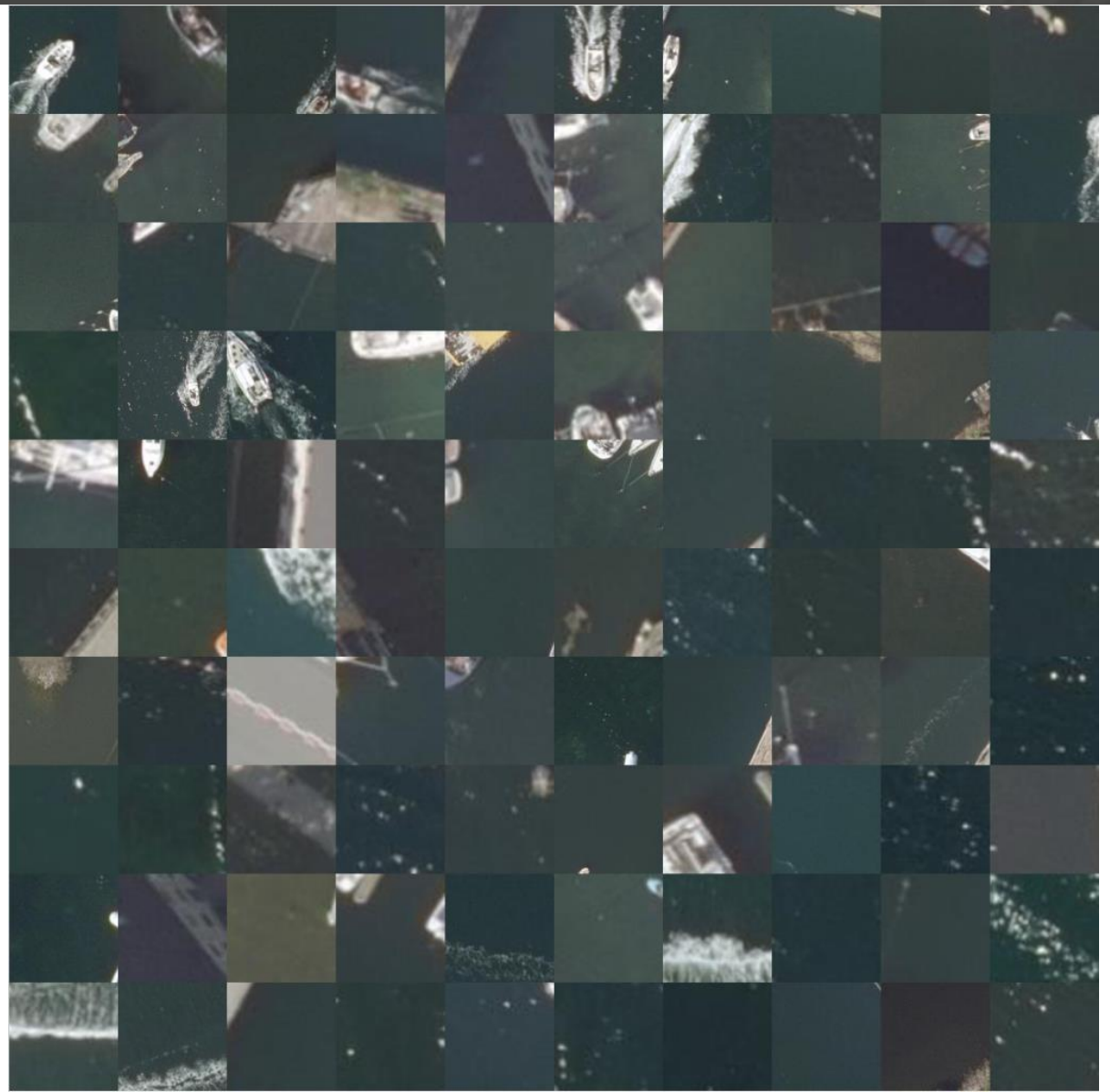
INTERACTIVE SEARCH REFINEMENT

- Project in cooperation with the City of Copenhagen

ALEXANDRA INSTITUTTET

INTERACTIVE SEARCH REFINEMENT

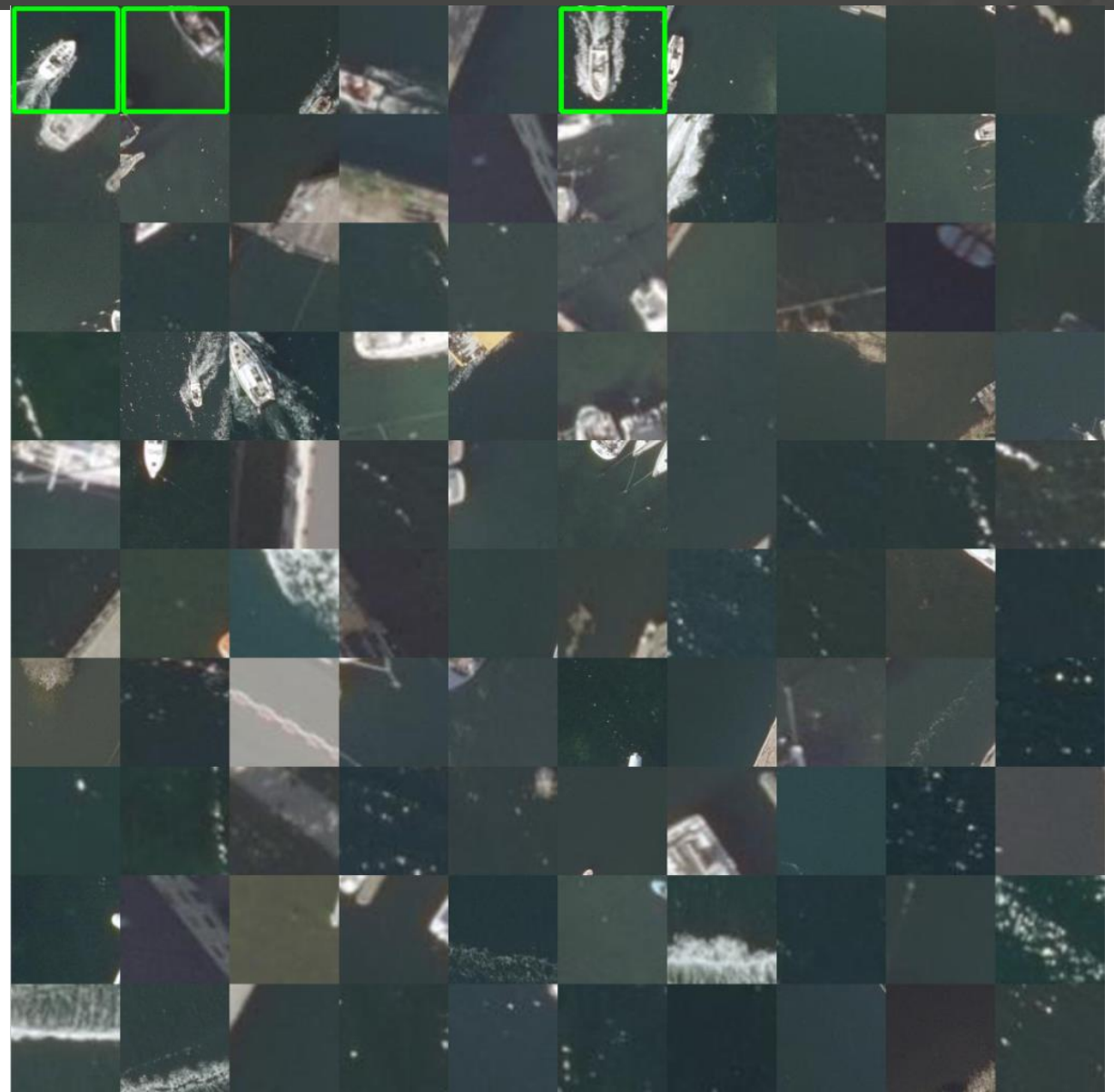
- Initial results



ALEXANDRA INSTITUTTET

INTERACTIVE SEARCH REFINEMENT

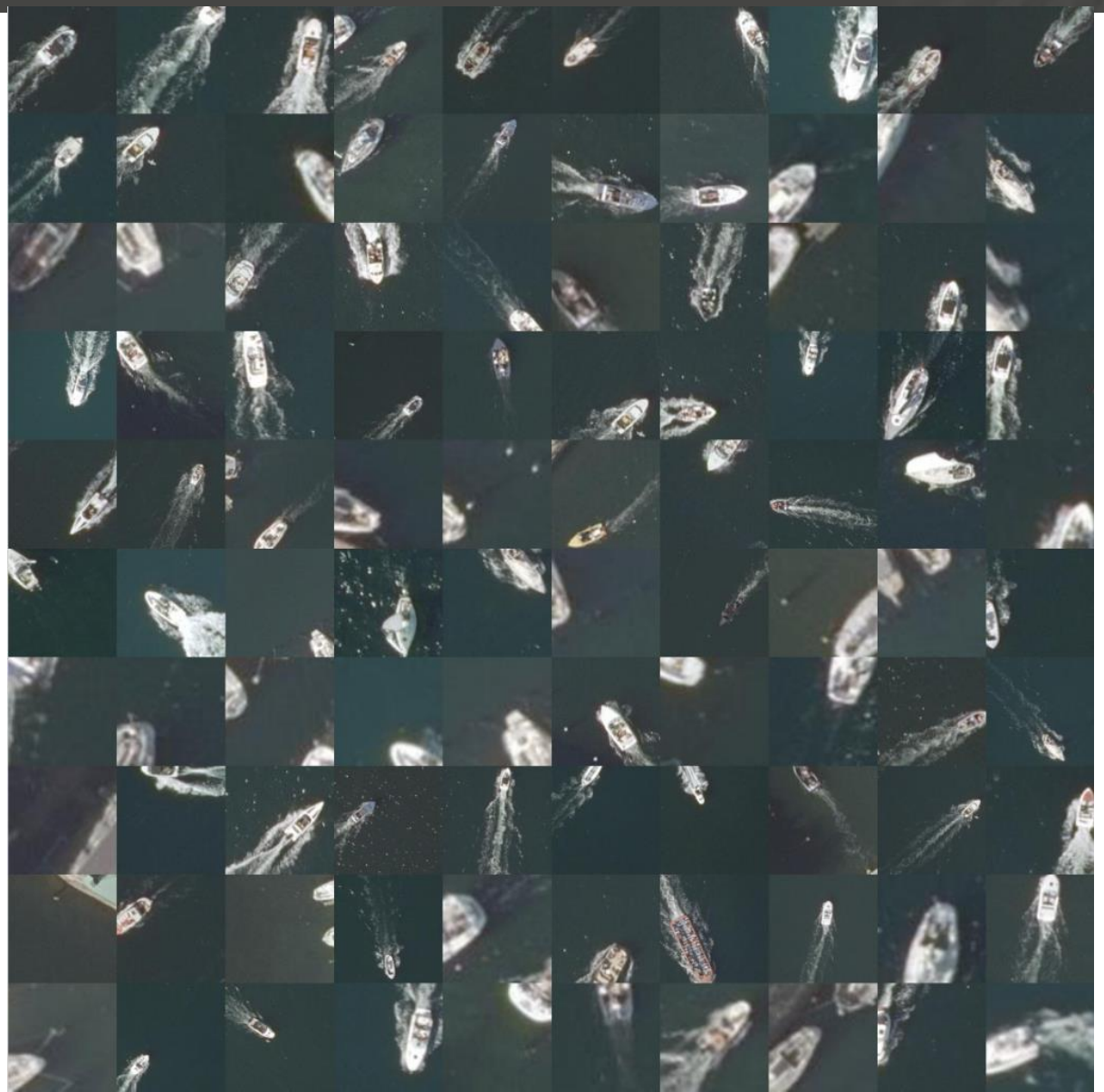
- Manual selection



ALEXANDRA INSTITUTTET

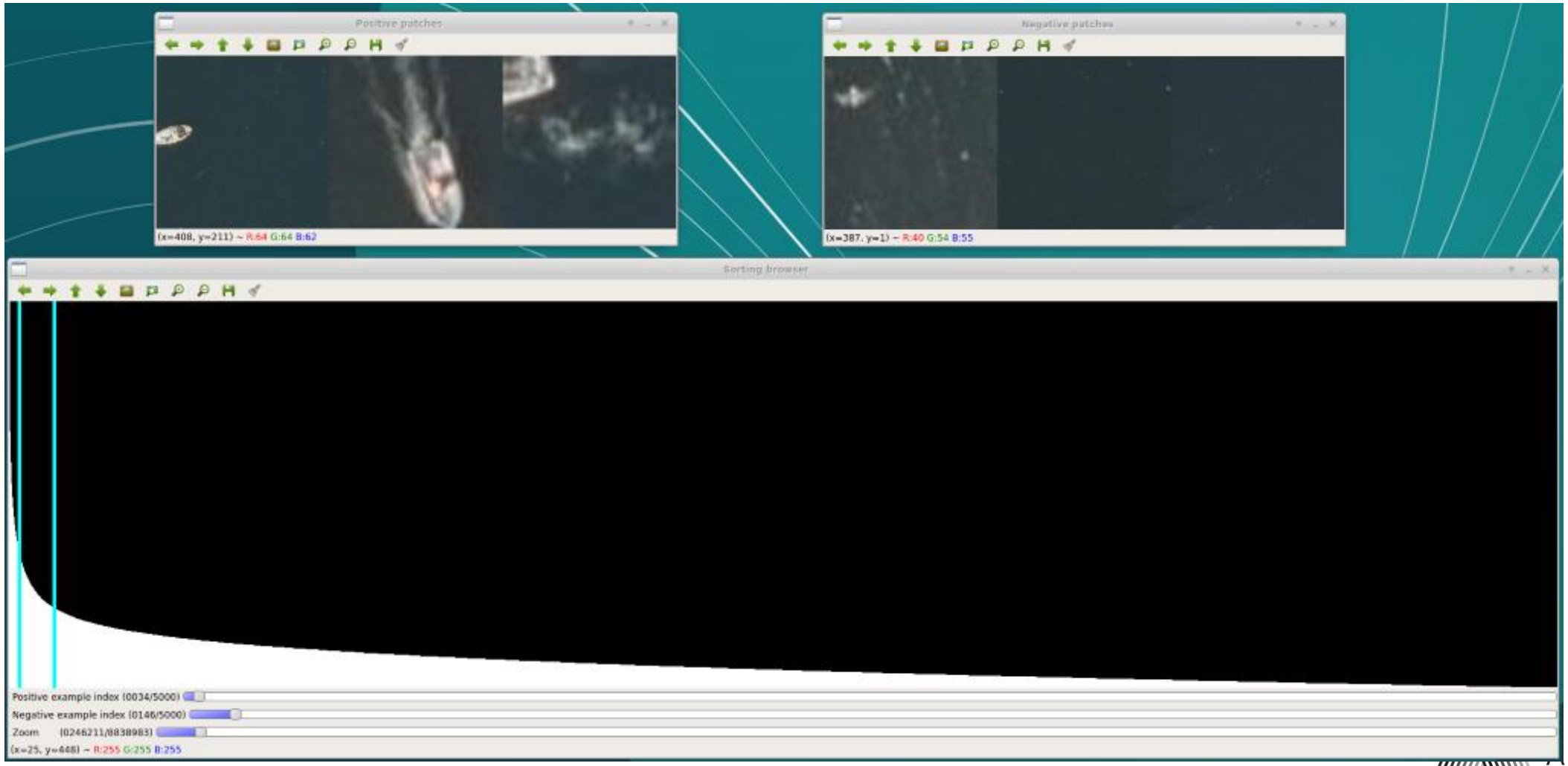
INTERACTIVE SEARCH REFINEMENT

- Refined results



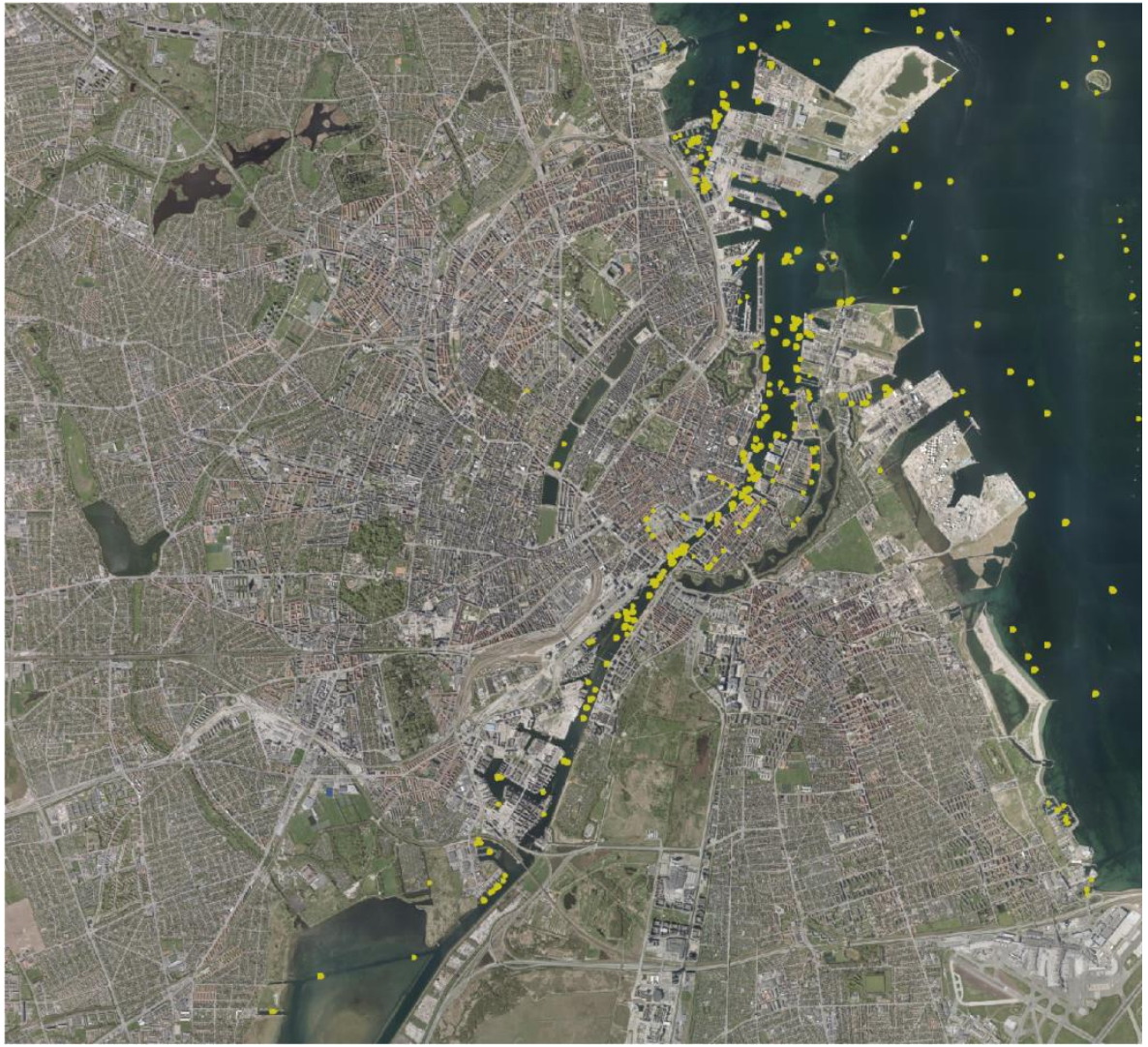
ALEXANDRA INSTITUTTET

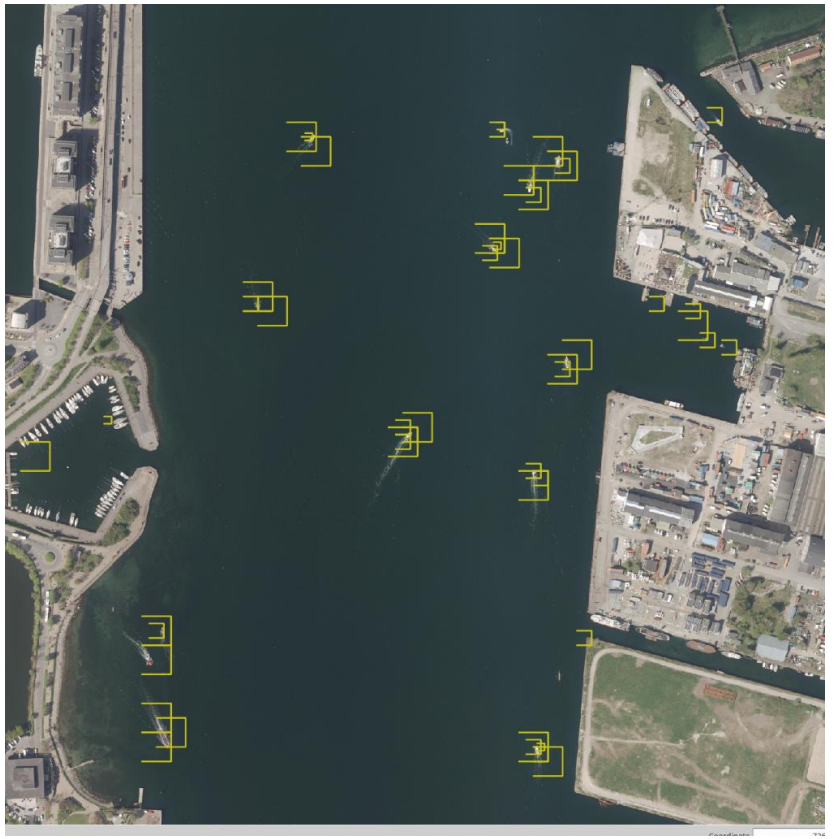
INTERACTIVE SEARCH REFINEMENT



ALEXANDRA INSTITUTTET

RESULTS: BOATS





ALEXANDRA INSTITUTTET

RESULTS: TREES

ALEXANDRA INSTITUTTET

REFLECTIONS

- Geovisual search gives promising new possiblities when it comes to performing fast searches for objects with very few query examples.
- Pretrained networks and cross-domain transfer learning
  - Does give encouraging results to use network for natural images
  - Would be better to base descriptor on network trained on orthophotos
- The method can be used to bootstrap training data generation for machine learning
- More experiments and use cases are needed to establish just how useful it is

ALEXANDRA INSTITUTTET

Tak for opmærksomheden!

ALEXANDRA INSTITUTTET