DETECTION OF BUILDINGS USING MASKININLÄRNING

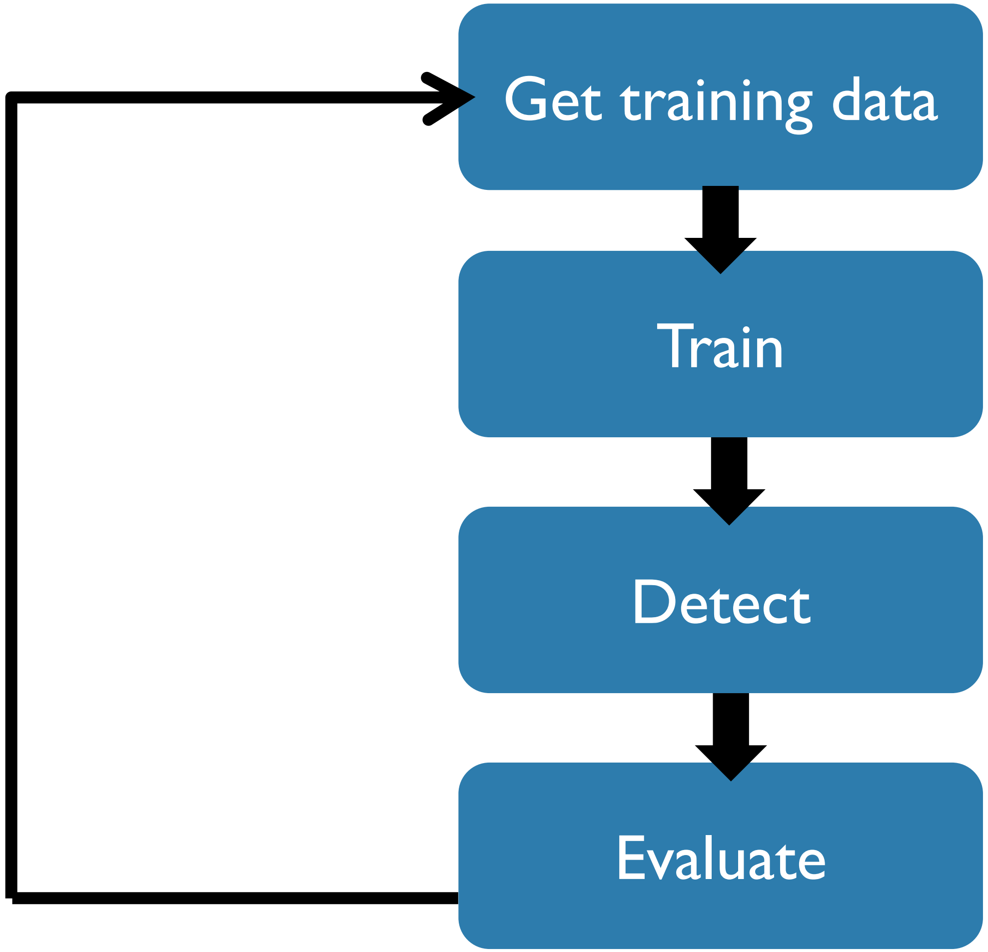
ERIK NILSSON 2024-04-18

LANTMÄTERIET

WHAT IS MACHINE LEARNING(ML)?

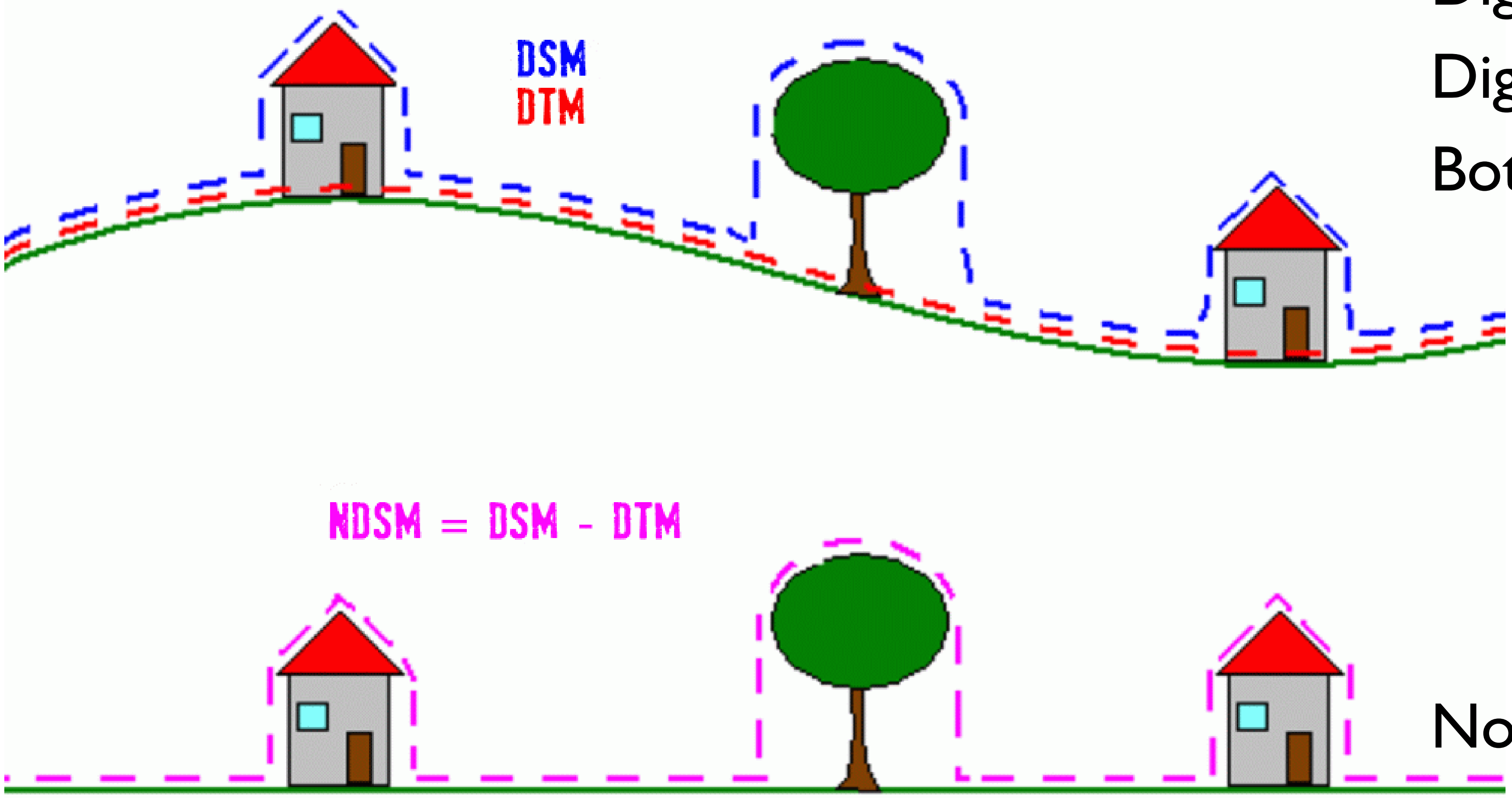
- In short: ML is a subdomain of Artificiell intelligence which handles methods for traning computers to detect and learn rules to solve a specific task.
- In our case it means that we train a model using the data we will use to detect something (orthophotos and a height model) and ”ground truth” data (building footprints) thats shows what we want our model to find (or detect).

DEVELOP A ML MODEL



ORTOPHOTO VS TRUE ORTOPHOTO

HEIGHT MODELS



Digital Surface Model (DSM)

Digital Terrain Model (DTM)

Both are ”height over sea”

Normalized surface model=
nDSM = DSM-DTM

”height over ground”

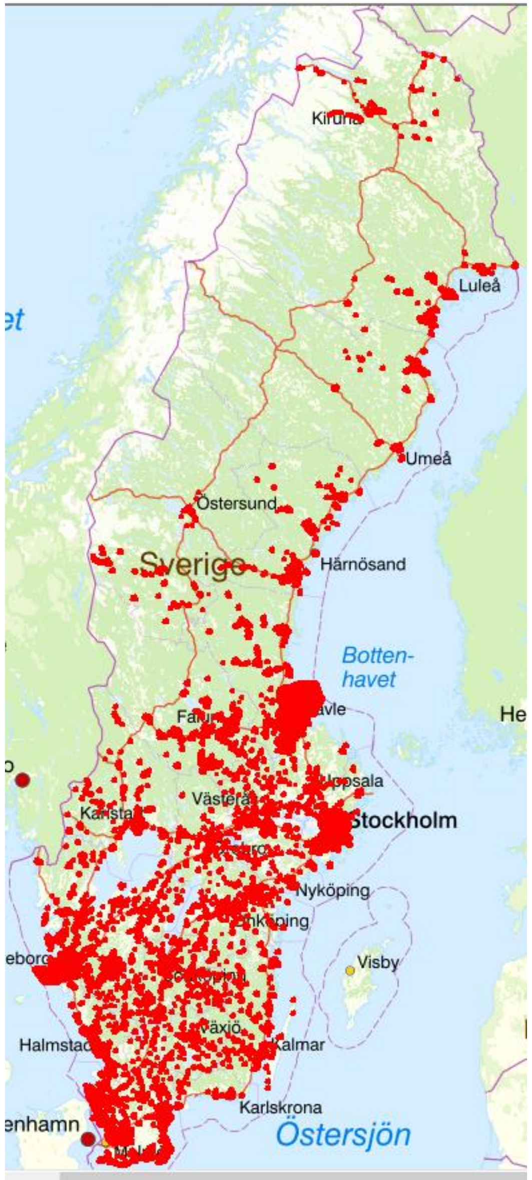
TRAINING DATA

Traning datat must have:

- High quality, both in geometry and completeness.
- Have diversity (many types of buildings, different types of terrains, different conditions in aerial photos (with/without leafes, dark/light images etc)).
- Qvantity (many examples to train on).

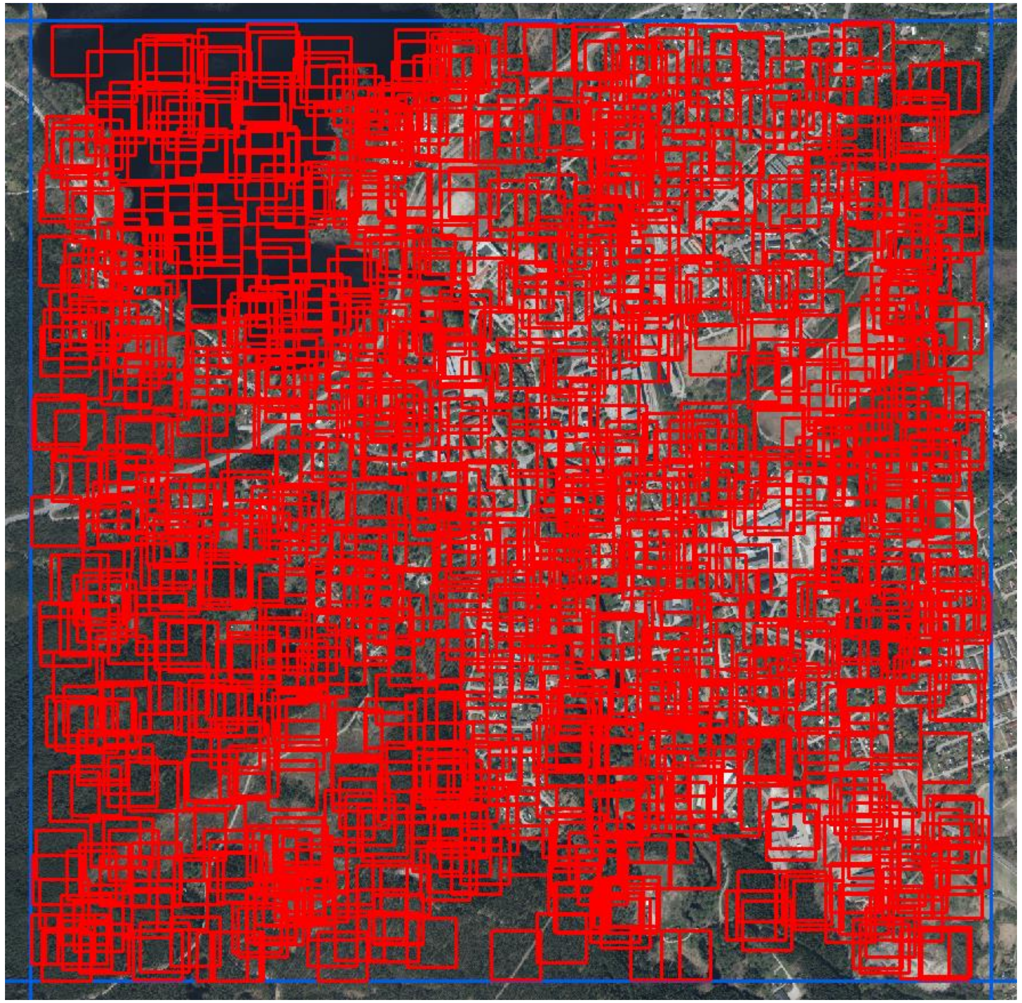
DATA FOR TRAINING

- The building data we can use is, in general, collected by the municipalities in the city centers and by Lantmäteriet in all other areas.
- This makes the quality and actuality varies a lot.
  Some municipalities report very small buildings and Lantmäteriet does not collect or make uppdates on buildings or changes less than (around) 15-20 m^2
- Lantmäteriets building data is, in general, not good enough to use for training.



TRAINING DATA

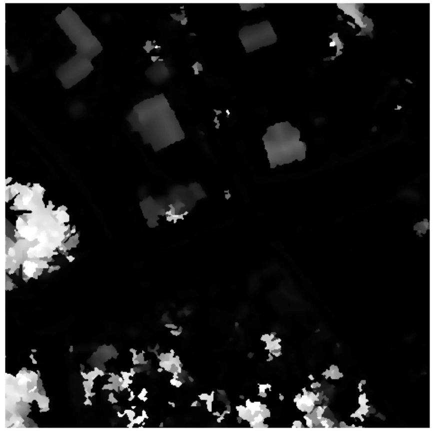
- We have randomly obtained tiles with data. These have been manually controlled and approved/nor approved.
- Tiles were also manually created to get training data for different types of cases like bridges, ports, camping areas etc. But also if our mappers have found strange results when evaluating our results.
- We have also mapped two areas 3*3 km very accurate and randomly created a lot of tiles from that areas.
- We have tried rotation and mirroring images to create more training data



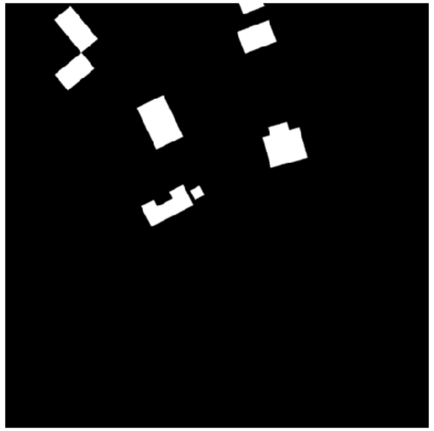
TRAINING

- We use tiles of size 512* 512 pixels. The resolution of the images we use is 0.25m så every tile covers an area of 128*128 m on the ground.
- We generate tif-images of all training data. This we wan use it for training even if new orthophotos are made or the buildings change in our databases.
- We have about 18000 tiles.
- The base data in all our training has been"True orthos" with R, G, B and IR- channels and rasterized buildings.
- We have tried severeal different height models like DSM, DSM and DTM and nDSM. We found that nDSM was the best in our tests.
- Traning a ML model needs good performance, a good GPU is required
- Our trainings takes about 4-5 days each
- When training takes such long time you need to be smart and patient when planning the trainings. Just change one thing at a time to be able to track what has effect on the results.

+

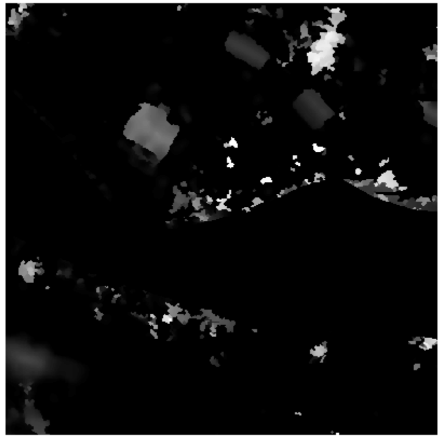


=



Training

+



=

Detection

+

+ML =

?

DETECT

Object found by ML

Ground truth

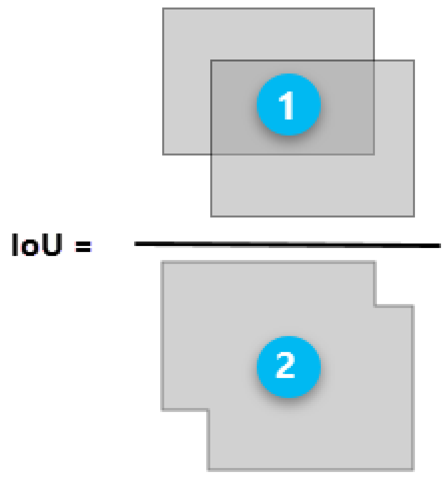
Area where ML and GT match

DATA FOR EVALUATION

- We have 35 different areas of size 2,5*2,5 km that we evaluate our models on. 5 of these areas have been controlled so that every building that are on the orthophos are mapped.
- We compare results using the ***Compute Accuracy For Object Detection*** method in ArcGIS Pro where buildings with area less than $2m^2$ are deleted.
- We also try to get a visual impression of the results.
- About 30 map sheets (10*10 km) have been evaluated by our mappers for testing and to give us feedback

STATISTICS

- The Intersection over Union (IoU) ratio is used as a threshold for determining whether a predicted outcome is a true positive or a false positive. The IoU ratio is the amount of overlap between the bounding box around a predicted object and the bounding box around the ground reference data.



Taken from https://pro.arcgis.com/en/pro-app/latest/tool-reference/image-analyst/compute-accuracy-for-object-detection.htm

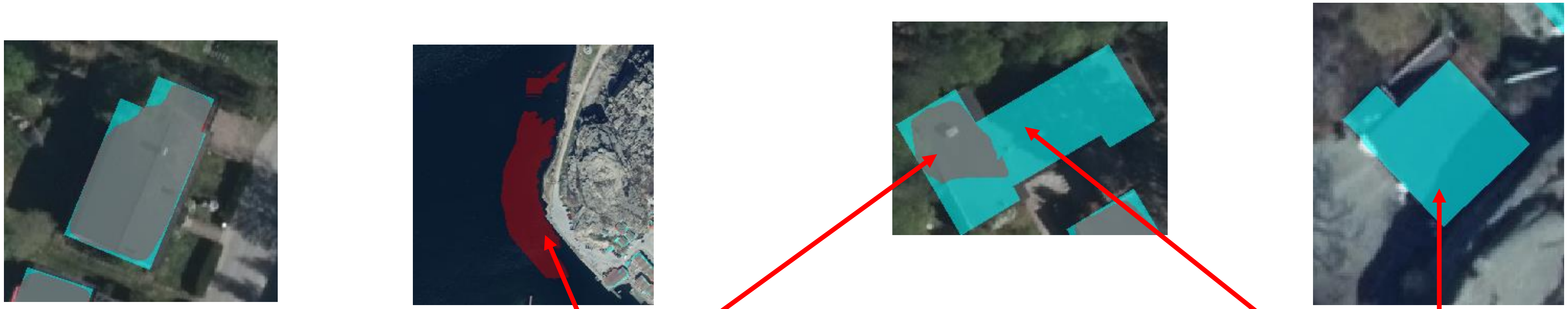
Object found by ML

Ground truth

Area where ML and GT match

STATISTICS

- We use a **IoU** threshold of **0,5**
- **True_Positive** — A detected object which has a IoU value >= 0,5
- **False_Positive** — A detected object which has a IOU value < 0,5 (this can also be an object found by the model where there should not be an object or an object found that not are in the ground truth data)
- **False_Negative** — An object that should have been detected but have not.



True positive False positive False negative

STATISTICS

- **Precision**—The ratio of the number of true positives to the total number of predictions.

  Precision=tp / (tp+fp)

  *The number of instances that are **relevant**, out of the **total instances the model retrieved.** (≈Classification accuracy)*
- **Recall**—The ratio of the number of true positives to the total number of positive predictions.

  Recall= tp / (tp+fn)

  *The number of instances which the model **correctly identified as relevant** out of the total **relevant** instances. (≈Completeness)*

Taken from https://towardsdatascience.com/precision-and-recall-88a3776c8007

$$F_1 = 2 * \frac{precision * recall}{precision + recall} = \frac{tp}{tp + \frac{1}{2} fp + fn}$$

EVALUATION

Blad	Precision	Recall	F1_Score	AP	True_Pos	False_Pos	False_Neg
T624_42_2500_Kullerod	0,8	0,78	0,79	1,02	117	30	33
T648_28_2525_Hunnebostrand	0,91	0,8	0,85	1,14	1925	182	473
T672_53_0025_Falun	0,88	0,72	0,79	1,23	2349	315	915
T684_57_0000_Ivarsmyran	0,61	0,66	0,64	0,93	35	22	18
T728_84_7500_Lovskar	0,83	0,6	0,7	1,38	464	96	311
Summa					4890	645	1750
Totalt	0,88	0,74	0,80				
Max	0,91	0,8	0,85	1,38			
Min	0,61	0,6	0,64	0,93			
Medel	0,81	0,71	0,75	1,14			
	0,12	0,08	0,08	0,18			

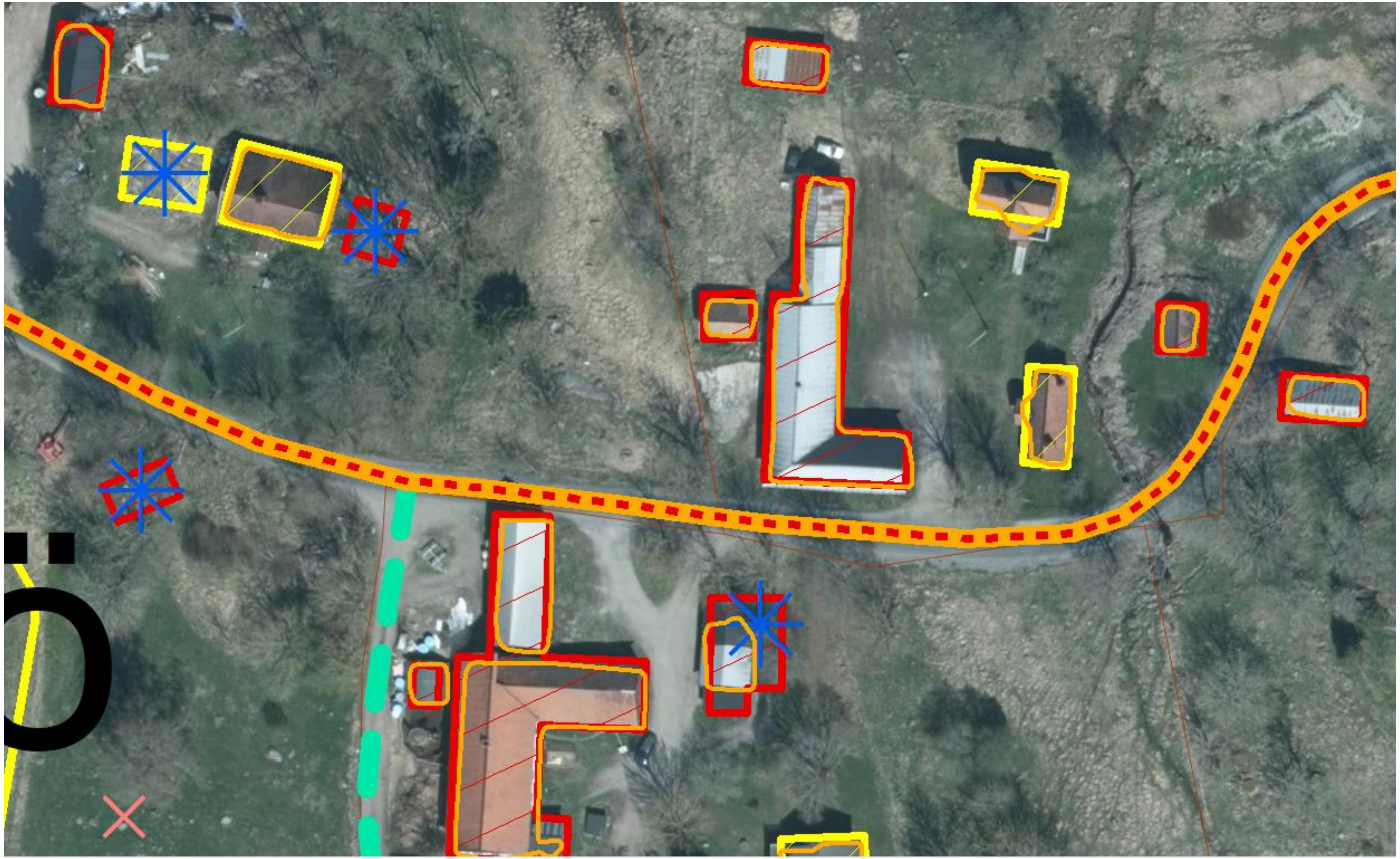
This means that 88% of the buildings found are real buildings and that we found 74% of the buildings that actuall

USAGE IN MAPPING

- Since the beginning of march we use buildings found by ML as a background layer for our topographic mappers
- The mapper decides if a change will be made in the database
- Tests indicates that this helps us find 1-2% more buildings than without support from ML buildings

ARCMAP – POLYGONS AND HIGHLIGTING DIFFERENCES

- ☑ ML_byggnad
  - ☑ ML_byggnad_polygon
  - ☐ ML_byggnad_punkt
  - ☑ ML_byggnad_diff



TECHNICAL STUFF

- NVIDIA A10 Tensor Core GPU on Vmware-platform
- Linux
- Python and ML framework PyTorch.
- FME for data management
- The technical environment is placed in a virtual ”container” which makes it easy to move to another computers.

WHAT WE HAVE LEARNED

- You cant create a good model from bad data
- We have used photogrametric surface models but think we would have better results with surface models from lidar-data. Lidar-data and ortho photos collected at the same time are high on our wish list.
- Takes a lot of time to train and evaluate models
- Be patient!
- Needs good performace on the hardware
- Make efforts to optimize the whole training process to make optimal use of the hardware

TACK! VI FINNS PÅ…

WEBBPLATS	www.lantmateriet.se
LINKEDIN	www.linkedin.com/company/lantmateriet
FACEBOOK	www.facebook.com/lantmateriet
INSTAGRAM	www.instagram.com/lantmateriet

KONTAKT	kundcenter@lm.se
TELEFON	0771-63 63 63

LANTMÄTERIET